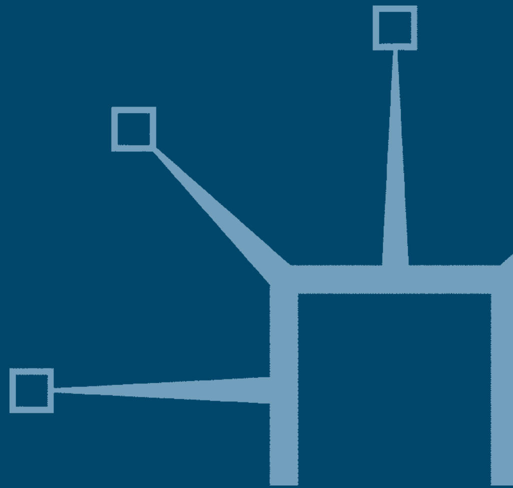
palgrave macmillan

Analytical Buddhism

The Two-Tiered Illusion of Self

Miri Albahari



Analytical Buddhism

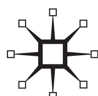
This page intentionally left blank

Analytical Buddhism

The Two-Tiered Illusion of Self

Miri Albahari

palgrave
macmillan



First published 2006 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN-13: 978–0–230–00712–3 hardback
ISBN-10: 0–230–00712–0 hardback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Albahari, Miri, 1968-
Analytical Buddhism : the two-tiered illusion of self / Miri Albahari.
p. cm.
Includes bibliographical references and index.
ISBN 0-230-00712-0 (cloth)
1. Self-Religious aspects—Buddhism. 2. Anatman. 3. Nirvana. 4. Buddhism—Doctrines.
I. Title.

BQ4262.A43 2006
294.3'422—dc22

2006047483

10 9 8 7 6 5 4 3 2 1
15 14 13 12 11 10 09 08 07 06

Printed and bound in Great Britain by
Antony Rowe Ltd. Chippenham and Eastbourne

Contents

Abbreviations from the Sutta Pitaka of the Pali Canon

AN	Aṇguttara Nikāya
DN	Dīgha Nikāya
Iti	Itivuttaka
MN	Majjhimā Nikāya
SN	Saṃyuttā; Nikāya
S	Sutta Nipāta
UD	Udāna
Miln	Milindapanha
Dhp	Dhammapada

Preface

For over two millennia, Eastern traditions such as Buddhism and Vedanta have claimed extraordinary capacities for the human mind. They have held that it is possible, through rigorous practices involving meditation, to attain *nibbāna* (Sanskrit, *nirvāna*), an exalted and irreversible mode of conscious existence. Some commonly reported characteristics of *nibbāna* include an end to all mental suffering. The capacity to feel negative emotion is completely burnt out. There is immense happiness, peace, equanimity and mental agility. Profound insight into the nature of mind and reality is discerned. Perfect virtue – in particular, compassion, generosity and loving kindness – is exemplified. It is not a rulebook morality; there is no *effort* to be virtuous. The pliable mind is spontaneous and present-centred, never lost in thoughts of past or future. There is ready access to objectless states of consciousness. A sense of identity as a separate autonomous self is absent, with no magnetic pull of 'me' or 'mine'.

The idea of *nibbāna* is not a museum piece for studies in ancient religion. People today devote their lives to its supposed attainment. In Western Australia, for instance, men and women shave their heads and don ochre or brown robes, renouncing their careers, entertainment, sex and relationships, to pursue what they believe will be far greater happiness. But have these Theravadin Buddhist practitioners got it right? Is *nibbāna* really possible? Or is it all just a fad, a chanting parade of ochre steeped in mystery without substance? And if something like *nibbāna* is possible, then how is it to be properly described? Are there real, not just apparent, convergences between descriptions of *nibbāna* in different Eastern traditions such as Buddhism and Vedanta? How much is religion and how much is reality?

I am deeply interested in the psychological possibility of *nibbāna*, both from a practical and philosophical perspective. The implications for Western philosophy (to say the least) would be vast, should its possibility become proven and recognised. I suspect that *nibbāna is* possible, but extremely difficult to attain. One possible (not foolproof) way to find this out is to join the ranks of the many who devote their lives to this 'ultimate' goal. Another approach could be to study advanced meditation practitioners in laboratory settings. My approach is somewhat more conservative. I seek to explore the possibility of *nibbāna* (in relation to Theravada Buddhism) from the angle of Western philosophy, presupposing no faith or esoteric belief. I treat it as a topic in the philosophy of mind. I want to see how far one can get towards arguing for the possibility of *nibbāna*, using analytic techniques and scientifically informed premises. This exercise in 'Analytical Buddhism' (a term I owe to Uriah Kriegel) is not meant as idle speculation; I believe it is needed

to focus any scientific study of the possibility of *nibbāna*, to help those in the laboratory know exactly what to look for. People are not born into *nibbāna*. If *nibbāna* is attainable, then one presumably has to train very hard to effect such a transformation. A transformation implies change from one state to another, so those studying the possibility of such a transformation will need to know what is supposed to change and what is not.

When setting out to write my Ph.D. thesis on the possibility of this nibbānic transformation, it soon became apparent what a mountainous project it would be. A battery of questions would have to be addressed before I could even get to the foothills of direct enquiry into *nibbāna*. This book is adapted from the thesis. I do not argue directly for the possibility of *nibbāna*, nor do I discuss the alleged ways to get there (e.g., meditation). But I do set out to arrive at the foothills. I hope that the philosophical journey there is interesting and coherent enough to support the kind of research that will move any future enquiry further up the 'mountain'.

So here is my approach. Above, I offered a brief depiction of *nibbāna* (and the list was by no means exhaustive). Investigating every one of these aspects – from their grounding in Eastern literature to their philosophical plausibility – is beyond the scope of a single book. My approach, therefore, is to take *one* central aspect of the *nibbāna* description – an aspect that has independent interest in the Western philosophy of mind – and explore it in its finer details. I seek to better understand, with regard to this central aspect, some of the earthy tangibles of what is supposed to change and what is not, during the course of this remarkable transformation. More precisely, I investigate whether this aspect is 'put together' in a such way that would potentially *allow* for the possibility of such a transformation. Should empirical or logical barriers prohibit this possibility, then the feasibility of *nibbāna* would be greatly diminished.

The aspect that I focus on revolves around the *self* – or the supposed lack of it. All the depictions of *nibbāna* allude to a transformation in which the practitioner starts out with a robust sense of individual self and then loses it upon the attainment of *nibbāna*. The losing of the self-sense is described as the divestment of a deep and powerful illusion. The self never existed in the first place. According to Buddhism, most of us mistakenly assume we are a conscious separate self, and until reaching *nibbāna* we are cocooned in this illusion. Upon attaining *nibbāna*, the cocoon is abandoned, but one does not emerge a zombie. In fact it is to the contrary. As we saw above, the consciousness of the nibbānic 'sage' is held to be greatly expansive in its native capacities. Such consciousness does not come into being; it is *always* there, but while enmeshed in the cocoon of illusory self it is unaware of its full potential. The transformation that I am interested to describe, thus moves from consciousness-*plus*-self-illusion to consciousness-*sans*-self-illusion.

My question is thus framed: is the ordinary self an illusion and, if so, is it constructed in such a way that could potentially allow for its dismantlement

(such that consciousness could *possibly* become liberated)? This question motivates the direction of the book. I argue that the self is an illusion contributed to by two strands or tiers. One tier is naturally unified consciousness – itself non-illusory. The other tier is grounded in a stream of desire-driven thoughts, emotions and perceptions. The content of these thoughts (and so forth) merge with native consciousness to create the impression of a conscious, unified, *separate* self. The illusion lies in the fact that while this self purports to think up the thoughts, the thoughts, in fact, help think up the self.

While that is the scope of my argument in the book, this model of self, if accurate, would green-light further enquiry into the possibility of *nibbāna*. The nibbānic transformation would, on such a model, consist in the gradual undoing of certain *types* of thought and emotion that are needed to preserve the sense of self. Their elimination would not erase consciousness but would leave it intact, uncovering any latent natural capacities. Investigation into such a possibility would have to face the hurdle of showing that it is consciousness *simpliciter* – and not consciousness with a sense of self – that carries even the most basic cognitive capacities that are needed to survive in the world. Some empirical studies in the West indicate that a sense of self may well, *contra* Buddhism, be required for autonomous survival.

The interest of this topic from a Western philosophical perspective should be apparent to those familiar with the vast literature on self, and in this book I engage with some of it. In particular I refer to the writings of Hume, James, Dennett, Flanagan and Damasio, eminent thinkers who have also denied the existence of self. There will be an important difference, however, between the way in which they construe the ontology of 'no-self' and the way in which I construe it. I say more about this difference in the Introduction, but, roughly speaking, these thinkers do not give consciousness (*sans* self-sense) the same import that I do. Features that I argue to be native to consciousness with their own independent reality (such as unity and uninterruptedness) are deemed an illusory facet of self by these thinkers. Should they be right, then it would be hard to see how (unified) *nibbāna* could be possible; all semblance of unity would for instance collapse if the illusion of self were to collapse. The role I give to consciousness – as a unified and non-illusory feeder to the self-illusion – is by contrast quite major and perhaps original to any Western theory that has denied the existence of the self.

The theory is not original to me, however. What I call 'the two-tiered illusion of self' prefigures (I argue) in the Canons of Theravada Buddhism (and more explicitly in Advaita Vedanta). The irony, therefore, is that common interpretation has Theravada Buddhism propounding an account of no-self that gives such a marginal role to consciousness that it is hard to see how *nibbāna*, by its own lights, could be possible. This 'Buddhist view' is indeed sometimes carted in to *endorse* the 'Humean' take on no-self. Because I am

interested in grounding my theory in what I believe to be the *actual* Buddhist reading – Buddhism deserves this much credit – I spend some time in early chapters extracting and justifying what I think is the correct construal of the Buddhist concept of no-self. If any philosopher should become fatigued at the prospect of a tedious wade through Buddhist exegesis, he or she should think again. While exegetical, the exercise is also eminently philosophical and, I hope, rewarding. I hope it will become more apparent that a seam of rich philosophical ideas lurks in the Buddhist Canon, waiting to be properly tapped (for instance their concepts of ownership and identification). To dismiss such ideas because Buddhism is commonly labelled a 'religion' is to epitomise the sort of dogma that can *give* religion a bad name.

Scholars of Buddhism may likewise be alarmed at my radical departure from the usual reading of the 'no-self' concept (compounded by the fact that my knowledge of Pali is limited!). Once again, I invite those readers to suspend their preconceptions. The process of (Theravada) Buddhist exegesis should not rest entirely upon the translation of each Pali phrase in its exactitude. While not doubting the value of careful scholarly translation, I hope to strengthen the idea that analytic philosophy can *also* contribute significantly to the understanding of Buddhist concepts like no-self (*anattā*). As part of my methodology (to offset any bias), I generally seek, when analysing contentious passages, to rely on translations by respected Buddhist scholars whose interpretation of 'no-self' (in its relation to consciousness) is at odds with my own.

This book is, of course, only one of a number of philosophical projects that attempt to bring Buddhist ideas to the West. I hope that its enquiry will serve the general project of bringing to life a part of a text that was canonised over two thousand years ago, such that it sheds light on theories of the self and consciousness in a contemporary philosophical setting. I hope that the arguments will be firm enough to serve as part of a foundation for future enquiry into the wider project that motivates this book: the genuine possibility of *nibbāna*.

From its inception as an idea for a Ph.D. thesis to its final incarnation as a book, I have been offered much support for which I am grateful. While I cannot list everyone, certain people and institutions made a particular contribution. My mother, Sonia, has been of tremendous support throughout the process, including her assistance in proofing the manuscript. John A. Baker (my supervisor at the University of Calgary) continued to provide his invaluable comments ('JABs') on successive drafts of the thesis during my time in Perth when I was not enrolled in the programme. My brothers Joseph and Ben were supportive throughout, with Joe writing a nifty programme that resolved a 'dire critical' situation with the Pali words. Michael P. Levine, at The University of Western Australia was encouraging from the start; it is to him that I owe the impetus of turning thesis to book. I heartily thank the scholar-monk Bhante Sujāto, whose incisive feedback compelled me to

rewrite Chapter 2 (although he may still disagree with my line of argument!). Kathleen Potter taught me how to use the speedy endnote programme, while Jane McKessar contributed to editing of the Introduction and Chapter 4. Speranza Dolgetta and Deanne Ivany were wonderful in helping me stay focused while finishing the thesis in Calgary, a time during which I profited from many discussions with Greg Janzen. Uriah Kriegel and Tim Bayne provided me with valuable comments on earlier drafts of Chapter 6, Uriah offering his friendship and advice during various stages of book production. I have also benefited from the comments of anonymous referees.

I am grateful to the (Theravada) Buddhist Society of Western Australia. Through attending years of talks and meditation retreats by their senior monastics, I learnt more about Buddhism than from any book. The Zen Group of Western Australia also helped in this capacity. I am indebted to the Philosophy Department at the University of Calgary for sponsoring my Ph.D., and to the Philosophy Department at The University of Western Australia (where I now work) for keeping me employed while working on the book this year as well as in earlier phases. I warmly acknowledge the Philosophy Department at The University of New South Wales, who aided my progress on the book by awarding me with a Visiting Research Associateship along with access to facilities while in Sydney last year. At each of these institutions, the administrative staff made my life much easier, in particular, Renilda van Aerden, Merlette Schnell, Lee Carter and Soon Ng.

I would like to dedicate this book to my late father, Michael, for his unwavering encouragement in my career as a philosopher. While he did not live to see the completion of my thesis, I am sure that he would have been delighted to know it became a book.

Introduction: The Two-Tiered Illusion of Self

In Western philosophy there exist many theories of consciousness and the self. While Eastern input into this debate is not yet mainstream, it is gaining momentum as Western interest in this tradition grows. As part of a wider project (mentioned in the Preface) this book will attempt to explicate and develop some Buddhist ideas on consciousness and the self into a substantial position in the philosophy of mind. The position will bestow illusory status to the self as a whole, but a non-illusory status to several features that are ascribed to the self, features that I argue are intrinsic to consciousness. What makes this account unconventional is that standard Western accounts of self-as-illusion, following in the tradition of Hume and James, usually consider most of these ascribed features (in particular, those I refer to as *unity* and *unbrokenness*) to have an illusory status in themselves; they are what *make* the self illusory.

The task of this book will first involve extracting from Buddhist primary sources what I think is the most accurate reading of the Buddhist position on consciousness and self (or *no-self*). The Buddhist school that informs my analysis is the *Theravada* tradition, whose teachings are based upon texts from the Pali Canon (in particular, discourses known as the 'suttas'). The Pali Canon is generally agreed by scholars to contain one of the earliest records of the Buddha's own teachings (Pali being the language probably spoken in central India during the historical Buddha's era of around 2500 years ago).[1] My depiction of the Buddhist position on consciousness and no-self will, based on inferences from these suttas, depart in significant ways from the depiction of consciousness and no-self that one is likely to encounter upon opening a reference book on Buddhism. A fair amount of space in early chapters (Chapters 1 to 3) will accordingly be devoted to arguing that my unorthodox reading of these concepts is, in fact, most in line with Buddhist suttas. Throughout this book, it should therefore be kept in mind that what for brevity goes under the label of 'Buddhism' is to be doubly qualified, first, as rooted in translations of Canonical texts that inform *Theravada* Buddhism and second, as based upon my own philosophical inferences drawn from various passages within those

texts. While there will, of course, be those who take exception to this reading, it should at the very least make clear that what has often been touted as 'the Buddhist view' on the status of self and consciousness is, actually, based on inference as well. The popularised view is not there explicitly in the words of the Buddha and neither is the view I espouse. The Buddha was not interested in asserting ontological doctrines of consciousness or the self; his *modus operandi* was not philosophy but praxis – to teach one how to end suffering. Gleaning semantic, ontological and metaphysical positions from Buddhist suttas is therefore likely to involve inference and speculation – and I hope that what I have to offer is more accurate than the party line.

Having extracted my reading of the Buddhist 'no-self' position (a reading that aligns Buddhism more closely with Advaita Vedanta than is usually acknowledged),[2] I focus on the task of arguing directly for a version of it in the philosophy of mind.

The first thing to be clear upon is that the notion of self talked about in this project is one that describes the ordinary garden-variety self that we – or at least most of us – reflexively assume we are. Roughly speaking, the self will, I argue, turn out to be a conscious subject that is a unified, happiness-seeking, unbrokenly persisting, ontologically distinct or *bounded* 'me' who is an owner of experiences, thinker of thoughts and agent of actions. On the proposed theory of 'no-self', then, a self of such a description will not actually exist, although most of us will, through our very mode of living, be reflexively *assuming* that we are such a self. In assuming that we are such a self – a self that turns out to not actually exist – we will therefore be in the grip of a deep-seated *illusion*. Now this illusion (and the general notion of *illusion* will be spelt out in some detail in the book) will, on the proposed theory, be 'fed into' by what I refer to as two 'tiers'. One 'tier' will be most actively responsible for the self's illusory status. It will involve input from the content of desire-driven thought and emotion, input that makes the self present as an *ontologically unique* or *bounded* entity. The other 'tier', while implicated in the self's overall illusory status, will not in itself be mentally constructed or illusory. It will involve input from 'witness-consciousness': that underlying factor which, I argue, makes the self seem as if it is *consciously apprehending* the world (regardless of the particular sensory-mental modality through which the world is apprehended).

In accordance with the Buddhist position (interpreted), it will be argued that witness-consciousness 'imports' certain features into the overall self-illusion, features that are not in themselves illusory. While these features will *seem* intrinsic to the bounded self, they will in fact, on the argued position, be intrinsic to the contributing witness-consciousness. The features (which receive a fuller description in relevant chapters) include what I refer to as 'elusiveness' (the subject systematically eluding its own attentive purview), 'synchronic unity' (discrete thoughts and perceptions being unified, at a time, to a single point of view) and 'unbrokenness' (a *stream* of discrete thoughts

and perceptions being unified, from one moment to the next, to a single unbroken point of view). What makes this theory of the self-illusion stand apart from other Western theoretical counterparts of the 'illusory self' – as exemplified in accounts by Hume, James, Dennett, Flanagan and Damasio – is that the Western accounts usually pin the self's illusory status on the unity and unbrokenness (with consciousness *qua* witness-consciousness ignored). Such theories tend to explain these features in terms of a 'bundle theory', where tricks of memory and imagination employ discrete thoughts and perceptions to create the illusory impression of a unified and uninterrupted self. An illusory impression of such unity and unbroken persistence is what, in these theories, makes the self an overall illusion. As it happens, this is the position most commonly ascribed to *Buddhism* by leading scholars in the field.

What I infer to be the more accurate Buddhist position on self-as-illusion, however, wholeheartedly rejects any such 'bundle theory' of unity and unbroken persistence. The trick that our mind plays on us, according to my reading of Buddhism, is one that rather results in the impression of being an *ontologically unique and bounded (separate)* entity that is unified, elusive, unbroken and so forth. It is our impression of being *separate and unique* that, on the Buddhist reading, primarily makes the self illusory. The unified, elusive, unbroken (etc.) awareness forms a non-illusory 'tier' which, when infused with a 'tier' of mentally constructed input, creates the impression of a bounded self. By virtue of its two contributing 'tiers', I describe this Buddhist-derived model of the self-illusion as a 'two-tiered illusion of self'.

The most vivid way to illustrate the two-tiered illusion of self is through the use of an analogy, which will be spelt out as the theory is developed in this project. The self that we assume we are, will as a matter of fact, have an ontological status analogous to that of a dreamt-of sound (e.g., a shrill voice) that has a non-dreamt component (e.g., the shrillness from an alarm). The dreamt-of sound, 'the shrill voice', is analogous to the self as construed on the Buddhist account; it is an overall illusion and construct. In a central sense, there is no such thing as the shrill voice and no such thing as the self (hence, 'no-self'). Nevertheless, the shrill ringing quality – derived from the alarm clock – is non-illusory; it is not dreamt up. There is such a thing as the shrill ringing quality. Similarly, while the self as a whole is construed as an illusion in Buddhism, various features ascribed to the self on this view – those pertaining to its witnessing, unified, elusive, unbroken and invariable character – are not considered illusory or constructed. These non-illusory aspects that most people reflexively and mistakenly attribute to the self correspond to and qualify an everyday witness-consciousness. The illusory self as depicted in Western theories, by contrast, is more akin to a dreamt-of voice whose shrill quality is entirely dreamt up, with no recognisable input from the world outside of the dream.

The book's goal, then, is to extrapolate and defend a Buddhist-based account of the two-tiered illusion of self along the lines just described. What follows is

a brief synopsis of how the argument will proceed in terms of each chapter. Chapter 1 sets the scene, where various important distinctions (e.g., 'subject and object', 'self' and 'sense of self') are introduced. It is a chapter intended to familiarise the reader with some basic Buddhist ideas (the theory of conditioned co-dependence, the five *khandhās*, the Four Noble Truths), so that it becomes easier to grasp what may be unfamiliar concepts in Chapters 2 and 3. Chapter 2 sets out to derive from Buddhist suttas the central notion of (unconditioned) witness-consciousness. Such consciousness, according to my reading of Buddhism, is both grounded in *nibbāna* and central to ordinary conscious states. The concept of witness-consciousness (when divorced from more contentious metaphysics) will be a key player in the project at large.

Chapter 3 then argues that this notion of witness-consciousness (as grounded in *nibbāna*) plays – or rather, ought to play – a major role in determining the famous Buddhist concept of *anattā* or no-self. If nibbānic consciousness contributes significantly to the *sense* of self, then it is logical to suppose that when Buddhism (implicitly) rejects the reality of a self, it is not necessarily rejecting the reality of every *feature* ascribed to the self. Specifically, it is not denying reality to those features of the self that are deemed to reflect, albeit distortedly, the character of the contributing nibbānic conciousness. Because they are distorted through their reflexive attribution to a self, such features would not, in the ordinary sense of self, parade as overtly 'nibbānic'. What stops them from seeming overtly nibbānic are deep-seated desire-driven tendencies to *identify with* and feel *ownership towards* various phenomena in the mind, body and world.

From Chapter 3 we thus derive, from the suttas, a detailed description of the self whose existence Buddhism rejects. It is a self that Buddhism supposes most people to have a sense of being. In Chapter 4, I spend some time arguing that most people do, indeed, assume themselves to be a *self* of this description. During the course of the argument, I further develop the key notions of *ownership* (of which there is more than one kind) and *identification*. While these notions have not been ignored in the philosophy of mind, they have, in certain respects, been under-analysed. I hope to contribute to their analysis in a way that draws upon my understanding of how the notions are used in Buddhism. This analysis of identification and ownership will prove crucial to my later theory of how the self-illusion arises.

In Chapter 5, we note that the existence of a self fitting this overall description has been denied by both Buddhism and some leading thinkers in the West. Given the differing metaphysical backgrounds from which the self's reality can be denied, we may wonder whether people always mean the same thing when claiming 'the self lacks reality'. The main purpose of Chapter 5 is to develop the concepts of *illusion* and *construct* such that they can serve as a common way to meaningfully deny reality to the self. Thus, when Buddhists claim that the self does not exist and when Hume and Dennett claim that the self does not exist, they may do so in a different way

and for different reasons. Yet, the basic content of their claim will be the same: an entity answering to the description of 'self' will not exist insofar as it is an illusion and a construct (in fact, we shall see that the self will be an illusion *if and only if* it is a construct).

In Buddhism (interpreted), the features of unity and unbrokenness are among those 'imported' into the sense of self through non-illusory witness-consciousness; so in Chapters 6 and 7, I attempt to establish this Buddhist claim. In Chapter 6, I seek to establish that a modified *concept* of witness-consciousness that I term 'awareness', (embedded in the concept of self), is to be specified with reference to *concepts* that depict those features (of the self) considered non-illusory in Buddhism. I claim that such specification does *not* extend to the concept of boundedness. With this background I then argue, in Chapter 7, that awareness, as specified by the features of unity, unbrokenness and so forth, is *in fact* not illusory or constructed. So *contra* Hume, James, Dennett *et al.*, awareness and its features *do* have the independent, non-illusory reality that Buddhism regards them as having. As a corollary to this, an interesting implication for the philosophy of mind is underscored in Chapters 6 and 7: awareness (whose concept has been given little attention in the philosophy of mind) turns out to be a necessary component of phenomenal consciousness. This is owed to the fact that those features which specify the nature of awareness (unity, elusiveness, unbrokenness and so forth) are each implicated in long-standing *puzzles* of phenomenal consciousness. So awareness, implicated in those features that make phenomenal consciousness puzzling, will also be a central feature of phenomenal consciousness.

In Chapter 8, I argue directly that the self is an illusion. I do this by attempting to show (with support enlisted from both Buddhist and neurobiological quarters) how the process of identification could plausibly and parsimoniously generate an illusion of self. The process will involve input from thoughts and emotions that seem to reflect an ongoing reflexive concern in the 'self's' welfare – an idea that is found in Damasio (1999). The content of these thoughts and emotions will, as Damasio supposes, reflexively depict *the self* as an *ontologically unique entity* who is engaging with the world, promoting its own well-being. But *contra* Damasio, I suggest that the first-person perspective, from which these thoughts and emotions are harboured, will not in itself be illusory but will involve the co-specified awareness (as demonstrated in Chapter 7). Through this perspective (aided by identification), the unconstructed awareness with its features of unity (etc.) will silently feed into the 'story line' of a self, creating the impression of a unified, unbroken ... bounded and separate self-entity that engages with the world. The two-tiered illusion of self, whose ground is argued for in Chapter 8, is described more formally in Chapter 9. In Glimpses Beyond, I outline how, as hinted at in the Preface, this theory can be utilised in a line of enquiry that may have far-reaching implications for the philosophy of mind.

1
Some Central Distinctions and the Four Noble Truths

In setting out to defend a Buddhist-derived theory of consciousness and no-self, we must first be familiar with some basic concepts and distinctions that will serve as building blocks for the discussions to come. This chapter sets out (a) to make the necessary basic distinctions and (b) to outline some key concepts found in Buddhist and related literature.

The distinction this chapter will first explain is one that is most fundamental to this project: that between 'subject' and 'object'. Suspending judgements on deeper ontology (to be addressed later in the book), I seek to propose a good working definition of this distinction, one that fits with pre-theoretical judgement informed by the first-person perspective. From this, I outline the Buddhist theory of 'conditioned co-dependence', followed by the distinction between 'self', 'sense of self' and 'person'. I then outline the core of Buddhist teaching, known as the 'Four Noble Truths'. The Four Noble Truths provide the context in which my interpretations of consciousness and no-self in Buddhism are to be understood in Chapters 2 and 3, as well as junctures at which to reiterate this project's line of argument.

1. Subject and object

In a short paper entitled 'Subject and Object', Mait Edey (1997) offers a useful explication of a fundamental distinction that we pre-theoretically make between 'subject' and 'object'.[1] Since the subject/object distinction will be pivotal to this project, it is worth quoting Edey at some length:

> For purposes of this discussion, let the term 'object' refer to anything anyone might be aware of or pay attention to. The term refers, then, not only to 'physical' objects, including whatever material processes, states, or conditions one might discriminate, but also to such 'mental' or immaterial entities or processes as pains, sensations, memories, images, dreams and daydreams, emotions, thoughts, plans, numbers, concepts, moods, desires and so on. Whatever we may think about their ontological or

> epistemological status in other respects, I hope we can agree at least that any of these may be objects of attention or pass in and out of awareness.
>
> Let the term 'subject' refer to I-who-am-aware, whatever opinion we may hold of what that 'I' may be. I hope that, no matter what various opinions we may hold, we can all also agree that: (1) I, subject, can be aware of some object; (2) I can focus awareness in attention; and (3) I can distinguish myself from the object I attend to.
>
> Here a crucial point should be emphasised. The distinction between subject and object, and our capacity to make that distinction, is prior to any particular opinion or theory about what either the subject or the object may be. Another way to make the point is to say that we make the same distinction, and make it the same way, regardless of what we may think we believe about the nature of the self or consciousness and their relation to the world. Yet another way to put it is to say that the distinction is not made on merely conceptual grounds. Any time you are aware of some object, or attend to some object, you won't have any trouble distinguishing it from yourself. That is, you're likely to know, immediately, without having to stop and think it over, or having to collect any evidence, which is you and which is the object. You can distinguish yourself as subject from any object whatsoever ('physical' or 'mental') any time you direct your attention to that object and realize that it is you who are aware and who pay attention, not the object. The real nature of the object and the real nature of the subject may be baffling mysteries to us, but these mysteries are no barrier whatever to knowing which is obviously which.
>
> (Edey, 1997, online)

This passage provides a good starting point for our analysis; *all* of us can distinguish ourselves as observing subjects from that which we observe. The distinction is cast by Edey as pre-theoretical, at least in the sense that, as explicated, the distinction does not require metaphysical commitments as to *what* subject and object actually are. Nor does it require knowledge of whether the distinction can be ultimately upheld, although it would *seem*, on the face of it, to resist collapse. Such questions – including whether subject and self are the same thing – will arise upon more in-depth analysis. However, our current task is to clarify some points made by Edey in a way that sharpens the distinction on a *pre-theoretical* level. We first focus on what is meant by 'subject'.

1.1. Subject

Edey's passage suggests two closely related aspects of what it means to be a subject, pertaining to (1) its *modus operandi* and (2) the perspective-creating locus for this *modus operandi* as it stands in relation to objects. The *modus operandi* of a subject seems, to put the point broadly, to be its realised capacity to observe, know, witness and be consciously aware. I shall use the

term 'witnessing' (or 'witness-consciousness') to cover all these modes of apprehension, but when I do so I am to be taken as talking only about phenomenal cases of such apprehension. By 'phenomenal' I mean that there is something it is like to be undergoing the apprehension (at this stage we put aside the question of whether such phenomenal character is contributed to by witnessing or is exhausted by objects that are witnessed). Minimally construed, witnessing can be described as the broadest mode of phenomenal apprehending, subsuming all species of conscious experiencing, perceiving, thinking and introspecting, whether these apprehendings are attentive or inattentive, human or non-human. Witnessing is, for example, involved in phenomenal perception that is visual, auditory, olfactory, proprioceptive, tactile or gustatory and in phenomenal cognitive apprehension such as thinking, introspecting, remembering, feeling emotions or imagining. If there is something it is like for a bat to navigate, then bat-navigation involves witnessing. Witnessing is the determinable under which any such mode is a determinate. It is not collapsible into any *one* of these modes; it is the 'percipient' or 'knowing' aspect common *to* all such modes, securing them membership in the set 'phenomenal apprehension'. The commonality of witnessing enables us, arguably, to tell *that* we are seeing, hearing, thinking and so forth; we do not *hear* that we hear, we *know* that we hear. *Attentive* witnessing involves a focusing of the conscious awareness (through whatever sense or mental modality) on some specific object of perception, thought or introspection (such as this writing); *inattentive* witnessing involves the peripheral conscious awareness we may have of surrounding objects such as the hum of a computer.

Now a subject does not present itself, in relation to observed objects, as something disembodied. It is not witnessing *sans* a point of view. As Thomas Nagel (1986) makes clear, whenever the world is viewed (with its objects), the viewpoint will always seem to be from *somewhere*. This view from somewhere relates to the second aspect of what it is to be a subject. What *seems* to delimit the parameters of witnessing – such that a *point of view* on the world is subtly discerned – is a psycho-physical instantiation in space and time: the framework of mind and body from which (and through which) witnessing seems to emanate. This extends not only to the viewpoint that delimits the parameters of witnessing within a waking, material world (with the physical body) but also to how things appear during dreams and suchlike; it is not as if the witnessing suddenly loses all sense of occupying a spatio-temporal perspective. *A subject is hence defined as witnessing as it presents from a psycho-physical (hence spatio-temporal) perspective*. For brevity I will sometimes refer to this as 'witnessing-from-a-perspective' or just 'witnessing perspective'.

Could the *modus operandi* of a subject, the witnessing, ever *come apart* from the specific perspective from which it seems to view the world? That is to say, could there be such a thing as pure witnessing with no determinate mode or perspectival locus – and hence no observed objects which would seem to

necessitate the perspective of a *subject* from which the objects are viewed? If witnessing could occur without a definite perspective (that is, without a subject), then the term 'subject', as I have defined it here, may refer to something that is not as psychologically basic as it pre-theoretically seems to be. In everyday experience, filled as it is with observable sensory or mental objects (whose observation demands that witnessing assumes a definite perspective), one is likely to get the impression that witnessing and a spatio-temporal perspective must always *go together* as a psychologically basic unit. It just so happens however, that Buddhism as I interpret it, regards witnessing *sans* a spatio-temporal perspective (and without any objects that are witnessed) to be a genuine psychological possibility that may (a) be realised through advanced meditative states and (b) is said (or so I shall argue) to reflect the real, unconditioned nature of witness-consciousness (without a subject as I have defined it). To avoid begging questions for or against this possibility, I refrain from formulating my definition of 'subject' or 'witnessing' in a manner that would either rule out or presuppose the psychological possibility of this scenario. While I presume the scenario of 'pure witnessing', as I shall call it, to be quite unimaginable (insofar as we would be psychologically unable to imagine what, if anything, it would be like), we can at least gain some *conceptual* handle on what the scenario of objectless, aperspectival witnessing might involve. It would imply witnessing without any input from the five senses: no sights, sounds, smells, tastes, tactile and proprioceptive sensations. There would furthermore be no cognitively sensed objects: no thoughts, concepts, directly observable emotions or desires. Without any objects witnessed, attentively or inattentively, including objects pertaining to one's bodily and mental state in space and/or time, there may well be no cues that could lend witnessing the implicit and ubiquitous impression of occupying a specific spatio-temporal perspective.

Why is it relevant to mention – at this early stage especially – this possible scenario of pure witnessing? What does it matter to our pre-theoretical definition of 'subject' whether its *modus operandi* of witnessing can or cannot occur without the perspectival confines necessary for a subject? It matters because *not* mentioning it may reinforce various assumptions about the subject – assumptions in line with ordinary object-filled experience – that Buddhist philosophy, as I understand it, would fundamentally reject. The assumption I have mainly in mind is that witnessing implies a subject; or to frame it in more detail, that the outlook of witnessing must always be confined to a particular (psycho-physical) perspective. Since this project aims to first *articulate* (before defending a version of) the Buddhist position on consciousness and no-self, it is important, from the outset, to set the stage for thinking about consciousness in a way that is accessible to – and does not prematurely rule out – the Buddhist position on consciousness. As I will suggest in Chapter 2, witness-consciousness, in Buddhism, is implicitly construed as something that is not (and does not necessarily seem to be)

limited *to* or *by* any psycho-physical (and hence spatio-temporal) perspective. Witness-consciousness is regarded as a metaphysically basic 'substance' (for want of a better word), completely unconditioned by what Kant would refer to as the parameters of time, space, quality and relation. At a later stage I will be saying more about what this involves; the aim for now is to make sure no questions are begged by supposing that whenever there is witnessing, there must be a delimiting *subject* that is witnessing. If the Buddhist position (as I interpret it) is correct, then there can be pure subjectless witnessing: that is, witnessing without a subject, viz., without the confines or appearance of a spatio-temporal perspective.

Allowing that the Buddhist position *might* be correct also calls into question a further common assumption: the assumption that witnessing must be categorically an *activity* or *state* of a subject. Such an assumption is at odds with (my understanding of) the Buddhist notion of witness-consciousness which is to be construed as metaphysically basic and *not* as a state or activity of anything else (and so perhaps as more like an individual). While this project will not argue for the reality of such metaphysically basic consciousness, it is hoped that what *is* argued for in this project will make more likely the scenario that witness-consciousness *could* be metaphysically basic in this way. While the surface grammar will make it sound as though I talk about witnessing as an activity or state, this surface grammar should not be taken to indicate any deep metaphysical reality, just as talk of the 'setting sun' does not imply that the sun literally sets.

1.2. Object

We turn now to the term *object*. In the following, I will refine Edey's stipulation. Edey stipulates the term *object* to refer to 'anything anyone might be aware of or pay attention to' – regardless of whether it be external or internal to the (witnessing) subject's mind. I will, more narrowly, define an *object* as *anything that a subject can, in principle, pay attention to*. I say *in principle* and not *in practice*, since there will be various things (e.g., the entire empirical universe, sub-atomic particles or bat-qualia[2]) where limitations of cognitive power (e.g., on the part of a human subject) will prevent them from being objects of a single attentive purview. While such items may in practice escape the focus of attention, they can in principle be attended to by a subject with the necessary cognitive powers, including an imaginary 'google-a-pus' with sensory organs to cover every atom of the universe. That is enough to render them objects on my definition.

This definition does not imply that something fails to be an object if inattentively witnessed, but it does mean that a subject must, in principle, be able to focus its attention on it. For example, two seconds ago, I (*qua* subject) was inattentively witnessing the hum of this computer. At this current point in time, however, I (*qua* subject) focus my attention on the sound, demonstrating that the hum of the computer is – and always was – an object. Thus,

anything that can in principle be attended to by a subject – and so viewed as distinctly separate from the subject *qua* witnessing – qualifies as an object on our analysis; the term is very general, cutting across a traditional distinction between individuals, events, properties and relations. This general designation of 'object' will be central to this project and neutral with regard to debates on representation. The term 'object' will, for example, remain neutral on the question of whether an observed object (e.g., a cat) reduces to a Lockean idea that partially resembles the feline object in terms of its primary but not secondary qualities, or whether the observed cat is a mind-independent object whose presentation is mediated via sensory peculiarities, or whether it is some phenomenon constructed by the Kantian categories of space and time in response to unknowable noumena, or whether … (and we can go on at some length). The point of central relevance is that the object is anything that can be attended to by a witnessing subject, whether *through* channels that are mental, visual, auditory, gustatory, olfactory, tactile, proprioceptive, or pertaining *to* those channels, to the very qualia *associated with* each sensory or mental modality.

While 'object' is defined in relation to a subject, we are not stipulating that an object be occurrently experienced by a subject in order for it to qualify as an object. It is enough that an object be the kind of thing that can *in principle* be attended to by a subject, whether that thing is 'internal' (such as a thought or mood) or 'external' (such as an apple tree). This includes the external world and its properties (occurrently observed or unobserved) as well as the world of thoughts, concepts, ideas, dreams, memories and pains – and hence it fits with Edey's intuitively drawn definition. Does it include the subject, viz., the witnessing-from-a-perspective? In other words, can a subject ever, in principle, be the attended-to *object* of an experience by a subject – whether the observed 'object–subject' be reflexively one's own witnessing perspective, or the witnessing perspective of another being? Regardless of the ultimate answer to this question, it should be noted that, for our purposes of providing a pre-theoretical criterion along which to draw the subject/object distinction, such a scenario should not, at the very least, *seem* possible. For if a subject could seem to attend either to itself or to another subject *as a subject*, then our criterion for distinguishing subject from object – a criterion that should uphold the distinction as *appearing* resistant to collapse – would be unconvincing. The subject would be just another object.

So how does the criterion fare? With regard to a subject attending to another subject (*as* a subject), it is fairly uncontroversial that even to have direct access to another subject's 'internal objects' such as its thoughts and pains is an impossibility – perhaps conceptually as well as empirically. So it seems even more unlikely – and I would say incoherent – that the very point of view from which another subject views the world could, *qua* that point of view, become the object of another subject's attentive purview

(a first-person witnessing perspective attending to another person's first-person perspective; the mind boggles!).[3] Can one's *own* witnessing-from-a-perspective (viz., the subject) seem like the reflexive object of one's attentive witnessing? Again, it would appear not. The witnessing perspective seems to systematically elude its own attentive observation; just as a human eye cannot directly see itself, the witnessing perspective cannot seem to attend to the very perspective from which witnessing seems to emanate. Nor does the witnessing *qua* witnessing seem observable as an object of attention. Whether all this turns out to be a benign point of logic (as Gilbert Ryle (1966, 19–42) suggests) or there is more to it (as will be argued later), it does not seem as if the subject can be its own reflexive object of attention.

Our criterion along which we draw the pre-theoretical distinction between subject and object thus nicely upholds its apparent resistance to collapse. It also serves to illustrate how Edey's criterion of objecthood as 'anything anyone might be aware of or pay attention to' is too loose. For while it seems impossible that a subject should be able to *pay attention* to itself, it does not seem impossible that a subject should be inattentively *aware of* itself. Edey's criterion, therefore, is in danger of bestowing a status of objecthood to the subject, thus collapsing the distinction in an all-too-easy manner.

Unless we encounter good reason to suppose that this pre-theoretically drawn or *apparent* distinction between subject and object does not hold, the distinction, as we have defined it, will remain axiomatic to this project. The actual *reality* of subjects and witnessing will be argued for much later in the project.

1.2.1. Objects and the Buddhist doctrine of conditioned co-dependence

I will now say more about the notion of objects in relation to Buddhist cosmology. In doing this, I introduce ideas that will aid in our later analysis of how to understand the concept expressed by the word *nibbāna* in Pali (*nirvāna* in Sanskrit). *Nibbāna* is depicted as being *unconditioned* – a notion that will make little sense unless we first understand what it means, in Buddhism, for something to be *conditioned*. That is the purpose of this section. Note that the exercise here is not to defend, but to elaborate Buddhist thought on this matter – although much of it should come across as fairly uncontentious.

In Buddhist cosmology, the world of objects – as I have specified the term 'object' – amounts to what is known as 'conditioned existence'. Each object within conditioned existence is characterised by a co-dependence upon other conditioned objects in a manner that is both *synchronic* and *diachronic*.[4] Synchronic co-dependence means that a given object cannot, at a given time, exist without a concurrently supporting network of other conditioned objects. 'Other conditioned objects' can include the internal structure and functionality of an object, in other words, objects that make up a more complex object. Our living bodies, for instance, depend upon the mutual functioning of innumerable internal objects such as organs, DNA, cells and

atoms. Supporting objects also occur outside of the conditioned object, our living bodies depending upon such factors as breathable air, temperature, gravity and sustenance. In Buddhist cosmology, there is no conditioned object – mental or physical – whose existence does not synchronically depend upon other specifiable conditions, such that the withdrawing of those conditions (themselves objects) would not bring about the destruction of that object. There is, in other words, no such thing as an indestructible object, whether this be an atomic particle or a depressed mood. At any moment in time, the synchronic existence of a conditioned object is made actual by the mutual manifesting of other favourable conditions. And each of these favourable conditions is itself dependent upon other conditions, suggesting a vast network of interconnected objects. That is partly what it means for an object to be 'conditioned'. (In C.B. Martin's (1996) system of dispositional ontology, this would be known as the 'mutual manifestation of reciprocal disposition partners').[5]

Along with synchronic co-dependency, Buddhism alludes to a diachronic co-dependency (or conditionality) of objects. This means that whether on a microscopic level (indirectly accessible to a human subject's perspective) or a psychological level (directly accessible to a human subject's perspective) each object over time is, due to other conditioned objects, undergoing creation, change or destruction in a non-random, law-like fashion. One micro-moment of objects will condition or influence the next micro-moment according to laws of nature that Buddhism says operate at a number of different levels (mechanical, biological, psychological and so forth).[6] While some of these laws (e.g., mechanical) apply more often to objects of the external physical world, Buddhism places particular emphasis on the train of objects that appear directly (even if inattentively) to a subject's perspective, and which, relative to this perspective, pass in and out of that subject's field of witnessing, or more colloquially, *awareness*. It is the set of objects as *occurrently* (not just dispositionally) viewed by a subject, whether attentively or inattentively. I shall for convenience refer to these objects as 'objects of awareness' or simply 'awareness-objects'.

Objects of awareness can include many items from the general set of objects such as the surrounding world, one's physical body, sensations, perceptions, thoughts, volitions and emotions. They comprise the totality of objects that are, either at one time or successively, experienced by a subject. Buddhism maintains that these objects, when divided into their smallest components, arise and pass away from our conscious awareness in much the same way that Hume (1739, 162) described when he tried to introspect the self: with an almost 'inconceivable rapidity'. The nature and flow of these objects in relation to one's awareness are of course subject to the spectrum of laws, such as mechanical/biological (stubbed toe leads to physical pain) or psychological (yearning for absence of pain leads to mental suffering). Buddhism maintains that there is not a single object of awareness – or indeed

object in general – be this thought, perception, situation, emotion or atom, whose existence has not been lawfully brought about or influenced by previous objects (where objects include states), whether from the preceding micro-moment or fifty years ago. There is no such thing as an object without a historical precedent: Buddhism does not accept that there can be something from nothing. Even the Big Bang is considered, in Buddhist cosmology, to be one of countless conditioned phases of the universe; conditioned existence is regarded as having no known beginning.[7]

1.2.2. Objects and the five khandhās

There are many ways of categorizing objects when the term 'object' is defined in the general sense I have introduced above. Buddhist literature suggests a particular categorization that, from their practice-oriented perspective, is most relevant (although not exclusive) to what I have termed 'awareness-objects', and which will be important in the discussions to come. The categories in question are, in Pali, called *khandhās* (Sanskrit: *skandhas*). These *khandhās*, which specify types under which objects (particularly awareness-objects) can be tokened, are: 'form' (*rūpā*: physical objects including bodies), 'feelings' (*vedanā*: pleasant, unpleasant and neutral sensory or cognitive qualia), 'perceptions' (*saññā*: recognition of sensory/mental objects and qualia), 'mental fabrications' (*saṇkhārā*: thoughts, desires, personality traits, intentions), and the object-oriented 'consciousness' (*viññāṇa*: awareness of objects specific to each sensory or mental modality).[8] In more detail, while 'form' pertains to anything physical, the body is alluded to in particular, with its sensory and cognitive receptors. 'Feeling' pertains to the sensation or raw feel of experiences that are associated with sensory or cognitive modalities, such as the redness of a rose, the aroma of coffee, the sadness of a memory or all of these combined. Feelings are further parsed into three useful categories: pleasant, unpleasant and neutral. Pleasant feelings are those associated with physically or mentally enjoyable experience; unpleasant feelings pertain to physical and mental discomfort or suffering; neutral feelings are all those which are neither pleasant nor unpleasant (such as the background hum of a computer). When it comes to explaining the 'Four Noble Truths' (later in this chapter) we will see that this parsing of feelings into the categories of 'pleasant', 'unpleasant' and 'neutral' is highly relevant. 'Perception' refers to the recognition of objects *as* particular objects, such as books, accidents, numbers and colours. It requires the functioning of memory, such that newly arisen input – mental or sensory – can be brought under a pre-existing concept. 'Mental fabrications' or 'formations' involve cognitive objects and events such as memories, volitions or intentions, thoughts, emotions, desires, attachments and character traits. Finally, 'consciousness' describes witnessing as it operates *through* the various sensory and mental modalities, taking various objects as a focus. It is witnessing as it is directed *towards* specific objects – attentively or

inattentively – and in this capacity it is co-dependent upon objects. So there is eye-consciousness, ear-consciousness, nose-consciousness, mind-consciousness, etc.

As one of the five *khandhās* of conditioned existence, this object-oriented notion of consciousness relates to specific modes of conscious witnessing rather than to the generic witness-consciousness that has been discussed so far. As such, the notion of *khandhā*-consciousness is one that I suggest straddles the distinction between witnessing subject and witnessed object (and so is not fully in the category of 'object'). The relation between witness-consciousness and *khandhā*-consciousness might usefully be thought of as analogous to the relation between white light and light that is refracted through a prism into its spectral colours. While each spectral colour partakes in the generic nature of white-light illumination, it is nevertheless distinguished in its hue from the other spectral colours. Similarly, while each moment of *khandhā*-consciousness partakes (I would suggest) in the generic nature of witnessing, it is distinguished from other moments of *khandhā*-consciousness by the object it takes at that moment. In this relational capacity, a unit of *khandhā*-consciousness is considered to be a highly impermanent phenomenon, in flux with the changing stream of awareness objects, even though it partakes in generic witness-consciousness that, as following chapters will suggest, does not intrinsically change with its objects. For pedagogical reasons, Theravadin Buddhism does not place emphasis on the concept of generic witness-consciousness; yet it will be argued in the following chapters that the concept of generic witness-consciousness is very much *assumed* in Theravadin Buddhism and, as such, is pivotal in understanding how the sense of self arises or can be disintegrated. For such reasons, it will be the concept of witness-consciousness rather than *khandhā*-consciousness that continues to feature in this book, the latter being mentioned in connection with the former.

1.2.3. The metaphysical status of objects in Buddhism

Having provided some overview of how objects (in my stipulated sense) are regarded in Buddhism, I finish with some general points about the metaphysical status pertaining to their purportedly conditioned, and hence impermanent nature. To do this, it will be useful to bring to bear on the exposition a fourfold categorisation of modes of conditioned existence. These modes can, I think, be found in Kant, but I want to use them without any commitment to Kant's theory about these modes. These modes in effect delineate key dimensions along which observable objects can be co-dependent; the four are space, time, quality and relation. To be conditioned by *space* is to have extension, with position and size. Physical objects such as the body will, on the Buddhist analysis, be conditioned by space. To be conditioned by *time* is to be either destructible and hence impermanent, or indestructible and existing forever *within* time (as opposed to existing outside of time). Since the five *khandhās* – whether mental or physical – are impermanent through their

synchronic and diachronic co-dependence, they will be conditioned by time. To be conditioned by *quality* means that an item will be positively definable through parameters that are empirically or conceptually delimited. Any quality that can be ascribed to an object will be definable in terms of those conceptual and empirical qualities that can be the focus of attentive purview. In the context of human observation, these will include observable empirical qualities that pertain to or can be known through the five sense modalities (discernible objects with colours, shapes, sounds, tastes, tactile and proprioceptive qualities) as well as qualities and concepts that pertain to mental deliverances such as emotions, thoughts and mathematical concepts. Obviously all tokenings of the *khandhās*, being observable in some way through sensory or mental modalities, will be conditioned by quality. To be conditioned by *relation* is to not exist independently but to rely upon other objects for existence. The conditioned synchronic and diachronic co-dependence of objects in relation to other objects (according to Buddhism) will ensure that no object can exist in isolation; hence all objects, on Buddhist cosmology, will be conditioned by relation.

The metaphysical status of objects in Buddhism will be alluded to again in Chapter 2 when we discuss (as best we can) the unconditioned nature of *nibbāna* and its conjectured relation to witness-consciousness.

2. Distinguishing self from sense of self

2.1. What is the difference between 'self' and 'sense of self'?

Before we can outline the notion of self relevant to Buddhism and this project at large, it is important to first get a clear grasp of the general distinction between 'self' and 'sense of self', since these categories are liable to be (and have been) confused. For the current purposes of explaining this distinction – as it will relate to the angle of this project – I will assume that 'self' means something like Descartes' *res cogitans*, a thinking subject with unbroken persistence in its identity over time. While this Cartesian notion of self will convey more than *just* what I have meant by 'subject', its central standing as a kind of *subject* will be shared with the fuller notion of self alluded to in this project, and so will have significant bearing on how we are to construe the sense of self. Now if 'self' refers to an enduring *res cogitans*, then it seems plausible to suppose that 'sense of self' will refer, at the very least, to a phenomenally felt, deep-seated *belief or assumption* that one *is* a self-entity of this description. For the purposes of the current discussion, I will assume that this deep-seated belief or assumption of being an enduring *res cogitans* is common among humankind. We should note that such a belief or assumption need not be reflectively obvious to the person who has it; just as one may be unable to articulate specific grammatical rules one has mastered, one may be relatively incognizant of the kind of entity one assumes one is (more on this soon).

The main reason for making a distinction between the self and the sense of self is that it enables us to articulate and explore a possibility that will be of central importance to Buddhism and this project in general. The possibility is this: while the sense of self widely exists, meaning that most people assume themselves to be a self-entity (a possibility that of course needs arguing), *the self itself does not exist* – meaning that there is no such entity that most people assume themselves to be (a possibility that will also need arguing). As it will transpire, the enduring *res cogitans* that illustrates our current example is not dissimilar in description to the self-entity that Buddhism rejects. We can at this stage guess that Buddhist cosmology will not allow for any distinct entity that persists unceasingly over time. (Note that my indexical usage of such terms as 'itself', 'ourselves', 'themselves', or 'oneself' is not meant to commit to the existence of a *self* in any controversial sense.)

Now, the rejection of such a self's existence is quite compatible with the fact that most people unwittingly assume the existence of this self by taking themselves (*qua* a subject) to be such an entity (viz., an enduring *res cognitans*). While the deep-seated and common assumption of self will on this scenario turn out to be a false one, it will nonetheless be true that the *assumption of self*, viz., the sense of self, will exist. Should this be the case, then the correct thing to say would be that while the sense of self is real (e.g., non-illusory), the self is not real (e.g., illusory).[9] I underscore this point so as to avoid falling into the easy trap of saying or thinking that the *sense* of self is unreal or illusory, when actually meaning that the *self* is unreal or illusory. Now should the self turn out, *contra* Buddhism, to be real, then it would simply mean that the sense of self is best explained by a self that is sensed. In this case, both the self and the sense of self would be real (e.g., non-illusory). Distinguishing the self from the sense of self, then, gives us logical room to investigate *whether* a possibly ubiquitous sense of self is explained by an actual self – as it would subjectively seem to be – or *whether*, as Buddhism maintains, it is in fact explained by cognitive conjurers that trick us into believing we are such a self.

It is helpful to give a parallel example that may be more familiar to some. Let us suppose that hard determinism is correct and that there is no such thing as libertarian free-will (such free-will is incidentally a feature commonly ascribed to the self that will star in later chapters). That is, we are supposing that it is not the case that, given a situation where we seem to exercise agency, we could have actually chosen (all other things being equal) to do otherwise. Every action is fully determined by factors of which none pertain to an agent's freedom to act otherwise. Libertarian free-will does not exist. Yet we can still entertain the idea that many people do harbour a deep-seated sense/belief/assumption/feeling that, given an identical situation, they could have chosen to act otherwise. This assumption of being a free agent, of having free-will, may well be real – despite the fact that free-will does not, on this scenario, exist. So while (on this given scenario) the sense

or assumption of free-will exists, libertarian free-will does not exist: the deep-seated assumption turns out to be a mistaken one. The hard determinists will attempt to explain the common belief in free-will not in terms of actual free-will – which would subjectively *seem* to explain it – but in terms of cognitive and psychological factors that do not include free-will. Similarly, while the sense of self may exist (as we are supposing in our above example), the self whose reality is assumed may not exist. If it does not exist, then the sense of self's reality will have to be explained by factors that do not include reference to *the self* in their ontology.

2.2. What is meant by 'sense' in 'sense of self'?

Now one may wonder at the choice of terms used to describe this deep subjective allegiance to the self's existence. While I have chosen the term 'sense' to be primary, my usage of other terms such as 'belief', 'assumption' and 'feeling' is meant to convey that the term 'sense' in this context is more complex than in some other contexts. The reason for allocating the word 'sense' as primary is that the turn of phrase 'sense of self' is already in vogue and, while lacking ideal precision, it captures the general gist very well. What, then, do we mean by 'sense' in this context? Let us distinguish it first from that associated with the five sensory organs, as put by the *Merriam-Webster OnLine Dictionary* (2006): 'specialized animal function or mechanism (as sight, hearing, smell, taste, or touch) basically involving a stimulus and a sense organ'. This is not the notion of sense we are concerned with, for the self, purporting to be a kind of subject rather than object, does not purport to be the kind of thing that could be detected via any of the five (object-tracking) sensory organs. The same dictionary offers, however, another definition that is more to the point: 'a definite but often vague awareness or impression <felt a *sense* of insecurity> <a *sense* of danger>'. One can have a sense of danger or insecurity without obvious input from a particular sense organ – which well suits the case of the self in question. The notion thus captures something more cognitive (as opposed to perceptual); a subjective or conscious impression of some sort. This notion of 'sense' is moreover not a success-term: to have a sense of X does not imply that X exists. For example, if one has a sense – or conscious impression – of danger, then there need not be danger that is sensed. This notion of sense, as a conscious impression, will thus apply well to the *self* whose existence may be under question.

As a kind of 'a definite but often vague awareness or impression', the term 'sense' as applied to 'self' has a further advantage. It manages to convey a subjective experience: that there is, in Nagel's famous phrase, *something it is like*, from the first-person perspective, to have or undergo a general conscious impression of X. While there may well be a host of non-conscious factors that help to generate the self-sense, it is nevertheless not akin to the wholly non-conscious machine-sensing that might be ascribed to a

thermostat. And it invites the question: just what underpins this impression – or sense – when it seems to indicate an X such as a self, danger or free-will? In the case of the self, we have already mentioned and questioned the most obvious option: that it is the *self* that underpins the sense of self. Given that 'sense' is not a success-term, however, other factors could contribute to it instead, factors that perhaps include what we have already designated as 'the subject', viz., a witnessing perspective that appears psycho-physically restricted. The virtue of this notion of 'sense', then, is that while conveying the idea of a first-person experience, it does not commit this experience to a particular genealogy but leaves the question open to enquiry.

I have also used the term 'belief' to portray something significant about the sense of self. I use this term broadly, to indicate that the sense of self conveys not only a subjective impression, but a propositional attitude with truth-bearing content that is deeply, and perhaps implicitly, assented to by the subject. The content of this belief, <this here is a self>, can thus be true or false, depending on whether there *is* such a self. If the self does not exist, then we can say that the assented-to belief is *false*; if the self does exist, then it will be described as a *true* belief. Now although we incorporate this doxastic dimension into what is meant by 'sense of self', we should be aware that such a belief taken on through harbouring a sense of self will differ markedly from the average belief that asserts the existence of an X. This difference will be mostly pinned on the fact that the self (that we are supposing we have a sense of being) purports, at its very core, to be a type of subject rather than object. To see what this amounts to, we can note that in ordinary cases of existence beliefs, the thing that is believed to exist is something that, given our way of drawing the apparent subject/object distinction, would form an object in relation to the subject who harbours, or appears to harbour, that belief. For example, when I say that I believe that this chair exists, the chair is a thing that, in relation to me as a subject, can be attended to in principle – it seems like an object that is separate to myself *qua* subject. Now although I may well seem (on our notion of the self as an enduring subject) to have an immediate sense of my own existence as a thinking, enduring thing that stands apart from other objects, and hence through this sense, a belief in my own existence as a self, the very nature of this self as being a kind of *subject* will dictate that I cannot focus my attention on the thing that I believe is myself (if I could, then that would render the self an object – and hence not a self!). Through its very nature as a (qualified) subject, the self could not, in other words, present itself as an object in relation to the subject that seems to uphold the belief. Rather than pertaining to a separate object, the self that is believed to exist will seem to integrate with the very subject that harbours a belief in the self's existence. In this way, any belief in the self's existence, insofar as it is harboured through a sense of self, will appear to be deeply reflexive. When harbouring a sense of self, then, the entity that we believe that we are will, minimally speaking, be the *subject that harbours that belief* in

the self's existence. Being this axiomatic, it will not be a belief that is easily revised if proved false (and hence will differ from most kind of beliefs that *are* easily revisable in light of contrary evidence).

With the sense of self there is hence, from the subject's perspective, no obviously *felt* gap between the believed-in item – the self – and the subject who holds the belief. ('Felt' in this context is another way of conveying 'sense', viz., a general conscious impression). Importantly, the subject believes that it is a self *not* by subscribing to some abstract intellectual treatise about the existence of a self – which would once again forge a felt gap between belief-holder and the state of affairs believed – but through deeply assuming the self's existence through its modes of thinking and living.[10] Hence my inclusion of the word 'assumption', suggesting something that is 'lived' and taken for granted. The nature and content of such an assumption, like mastered grammatical rules, will not be introspectively obvious to the subject that harbours it. The assumed self will not after all come across as an attentively observable *object*. Since a widespread assumption of self will colour the very way in which we as subjects approach and think about the world, it will be difficult to gain enough cognitive distance with which to introspectively discern its supposed content. We will have to resort also to other methods to get a picture of this self that we (may) have a sense of being – a task to which Chapter 4 is dedicated. Like an assumption of free-will, any subscription to the self's existence will be implicated in the very way that we approach the world: in our patterns of behaviour, thoughts, desires, emotions and motivations. Someone who assumes she is a free-willed agent, for instance, need not reflectively believe that she is free or that she harbours a belief of this kind. Yet her belief *qua* assumption can arguably be betrayed through becoming racked with guilt at making a bad decision. Unless she assumed she *could* have freely chosen otherwise (all other things being equal) she would not, it would seem, feel guilty. Similarly, someone who harbours a belief that he is a self, viz., an enduring thinking entity, need not reflectively know that he identifies as such an entity. His assumption can be displayed in an outbreak of nerves before having to give a speech. The nerves arguably suggest an implicit belief on the part of the apprehensive subject: a reflexive belief that the very same – and hence *enduring* – entity will give the speech and possibly make a fool of himself. It is a belief to which he need not, and probably will not, overtly subscribe.

The fact that the content of such a belief must be figured out indirectly is, indeed, a further factor that renders unusual the belief pertaining to the sense of self. Normally, a simple act of introspection will be enough to lay bare the content of a harboured belief or assumption, for example, the belief that Sydney has a famous opera house. A similar act of introspection will not, by contrast, reveal such an assumption as that of being an enduring, thinking, free-willed … self-entity.

2.3. The Buddhist perspective on the sense of self

I have mentioned that from a Buddhist perspective, the distinction between self and sense of self will be a crucial one to make. When reflecting on the kind of thing that a sense of self must be in relation to a self, a constraint is immediately placed upon the kind of self-entity that will concern the eminently practical orientation of Buddhism. It must be an entity that we, as ordinary subjects, could plausibly have a *sense* of being. The notion of self that primarily concerns Buddhism will thus not be an obscure metaphysical concept that is the exclusive domain of some intellectuals educated in traditional Western ideas. Since a fear of death or non-existence is so common, an immortal soul-entity is for instance unlikely to be the kind of thing that *most* people *could* have a sense of being (and scholars sometimes tout the self alluded to in Buddhism as an entity of precisely this kind). The notion of self that Buddhism primarily alludes to – a self that most subjects are purported to buy into – will rather depict a very ordinary kind of thing: an entity whose existence is reflexively subscribed to by almost every person on the planet. In Chapter 3, I will attempt to discern from Buddhist writings just what this supposed self amounts to; in Chapter 4, I aim to show that this kind of self-entity is indeed something that most people – Buddhist or not – have a sense of being. In Chapter 5, I discuss what it means to deny existence to this self and in later chapters I argue there is in fact no such thing as this self.

This book, therefore, does not take a revisionary approach to the nature of self – one that claims us to be wrong about the kind of self we assume we are (before offering an alternative account of what the self *really* is). Rather it seeks, with help from Buddhist analysis, to sketch a descriptive portrait of that *very thing* that we assume we are – via the *sense* of self – before investigating its ontological status. For this reason, a measure of privilege will be given to the first-person perspective (how things seem to the subject – even if only implicitly) when determining the supposed nature of self.

3. Defining 'person'

In a context where the existence of ordinary selves is being questioned, we need a relatively neutral term that enables us to refer to people or individuals in a manner that does not presuppose too much about selves or their sense. The term 'subject', while less controversial than that of 'self', is not ideally neutral. I noted earlier that a subject is not, *qua* subject, directly observable from either a first- or third-person perspective, and this may foster some doubt as to whether subjects actually exist. Now when *describing* the sense of self, reference to a subject will be unavoidable, since it will be a central component of the (assumed) self – and so if subjects do not exist then neither will selves. It will be useful, therefore, to have a term that does not presuppose anything too central about the ontology of selves. The term 'person' will be reserved for this purpose. 'Person' will thus, in this project, be a term that

designates a psycho-physical entity to whom selves are normally (even if mistakenly) ascribed. A person will appear to have the capacity for subjective witnessing, along with the kind of sensory and cognitive faculties that are normally (even if not necessarily) suggestive of a sense of self.

4. The essence of Buddhist teaching: The Four Noble Truths

I will now outline the core ideas in Buddhism, known as the 'Four Noble Truths'. This outline – which will make reference to such phenomena as *kamma* and rebirth (to be described) – may seem out of place in an investigation of the kind of issues I will be examining. However, it will provide the necessary background against which to understand my interpretation of the Buddhist position on consciousness and no-self in Chapters 2 and 3. When I use the turn of phrase 'what the Buddha said' or 'the Buddha's teachings' I mean to refer to what the historical Buddha was alleged to have taught, without taking a stance on whether he literally said those things. I will also allude to the 'Noble Truths' with capital letters to indicate that they are referred to by that name in Buddhism, rather than to presuppose that they are, indeed, true.

I have mentioned that Buddhism is an essentially practical body of teaching. By this, I mean that the focus of Buddhist teaching is not on expounding complex theoretical doctrine for the sake of intellectual curiosity, but on eliminating suffering. Any metaphysical-sounding claim such as that pertaining to the conditioned co-dependence of objects, or to *kamma*, rebirth and the unconditioned *nibbāna*, is only there in context of the Buddha's practically oriented teaching on how to eliminate suffering. The essence and emphasis of the Buddha's teachings on suffering, etc., is captured in this very first discourse that was given by the Buddha on the Four Noble Truths:[11]

> Now this, monks, is the noble truth of suffering: Birth is suffering, aging is suffering, death is suffering; sorrow, lamentation, pain & despair are suffering; association with the unbeloved is suffering, separation from the loved is suffering, not getting what is wanted is suffering. In short, the five clinging-aggregates are suffering.
>
> And this, monks, is the noble truth of the origination of suffering: the craving that makes for further becoming – accompanied by passion & delight, relishing now here & now there – i.e., craving for sensual pleasure, craving for becoming, craving for non-becoming.
>
> And this, monks, is the noble truth of the cessation of suffering: the remainderless fading & cessation, renunciation, relinquishment, release, & letting go of that very craving.
>
> And this, monks, is the noble truth of the way of practice leading to the cessation of suffering: precisely this Noble Eightfold Path — right view,

right resolve, right speech, right action, right livelihood, right effort, right mindfulness, right concentration.

(SN LVI.11)[12]

Rephrased, the Four Noble Truths – using key Pali terms – are as follows:

1. *Dukkhā* (approximately, suffering, stress, unsatisfactoriness, unfulfilment) exists.
2. The origin of *dukkhā* is *taṇhā* (approximately, craving, thirst, attachment, emotional investment).
3. The cessation of *dukkhā* lies in the cessation of *taṇhā*. The cessation of *dukkhā* involves *nibbāna*.
4. There is a path to the cessation of *dukkhā* (and hence to *nibbāna*): the Noble Eightfold Path, which involves the practice of insight-wisdom (*paññā*), meditation (*samādhi*) and virtue (*silā*).

Throughout this book, I will primarily use the Pali term '*dukkhā*' (rhyming with 'book-ah') for what has most commonly been translated as 'suffering' and the Pali term '*taṇhā*' for what has most commonly been translated as 'craving'. While '*nibbāna*' has no common translation, it has sometimes been alluded to as 'enlightenment' or 'awakening'. The reason for preserving these Pali terms is that there is no single-word English equivalent that captures their meaning. While my preference is to avoid non-English terminology when possible, the common translations of 'suffering' and 'craving' can be easily misunderstood; the English words are too crude and extreme to properly convey the right nuances. For instance, John A. Baker (in conversation) has rightly remarked that 'craving' in modern English carries connotations of rabid nicotine withdrawal symptoms! In outlining the Four Noble Truths, we will see that the Pali terms are rather more subtle and layered than this. Note that during this outline, the same remarks will apply as to when I outlined the Buddhist notion of conditioned co-dependency: we can assume that the Four Noble Truths are being *explicated* rather than *defended*. It is important to keep this in mind since, when the context demands it, wider Buddhist cosmology will be invoked.[13]

The First Noble Truth, then, is that *dukkhā* exists. *Dukkhā* alludes to the *khandhā* of unpleasant feeling, whether mental or physical, overt or covert. On a physical level, it includes all overt suffering such as that associated with sickness, injury and nicotine withdrawal. On a mental level, it includes any overt suffering from pangs of guilt and disappointment, to major depression, to not wanting a physical pain that is present, to craving a cigarette that is not present. But along with that, *dukkhā* pertains to the more covert unpleasantness that can occur while one is enjoying oneself. An example of this is needing to use the bathroom while listening to a wonderful concert. Even if the discomfort is not attended to but only

peripherally witnessed, it is still *dukkhā*, a *khandhā* of unpleasant feeling. There is also a kind of *dukkhā* referred to in Buddhism as the *dukkhā* of change or impermanence. A 'this won't last' thread of background anxiety, for example, can mar an otherwise enjoyable time. A further example of *dukkhā* during pleasure is the feeling that although one is enjoying an experience (such as a mountain view) one is somehow missing out on the amount of enjoyment that one feels one *ought* to be gaining from it. Luang Por Sumedho (1992, 23) once remarked that looking at the mountains in Switzerland was almost painful, as if his mind could not get – and hold onto – its rightful fill of hedons! This *dukkhā* is a feeling of unfulfilment, a sensing that one is not as full of happiness and joy as one might be. A related but more active kind of *dukkhā*, also very common, is what I shall term 'doer-*dukkhā*'. Doer-*dukkhā* pertains to anxiety at the thought of not getting the best possible outcome from one's actions (e.g., as in the thought: what happens if I do not win this competition?). *Dukkhā* thus pertains not only to the out-and-out suffering of an experience that is on the whole mentally or physically painful, but to the threads of unease, physical pain, dissatisfaction, anxiety or unfulfilment that can occur in an experience that is on the whole enjoyable.[14]

As well as pertaining directly to physical and mental states of persons, *dukkhā* is commonly used as a term in Buddhism to apply to all of conditioned existence, hence tokenings of the five *khandhās*: physical form, feeling, perception, mental formations and object-oriented consciousness. But what does it mean to depict the world and all its objects as unsatisfactory or suffering? My computer sure gives me *dukkhā*, but how can it *be dukkhā*? Indeed, how can the *khandhā* of pleasant feeling be *dukkhā*? The puzzle can be resolved through interpreting such ascriptions not as intrinsic, but as relational, which means that any object, if clung to as a source of happiness, has the potential to elicit *dukkhā* in that person. Put another way, it means that no object has the potential to elicit in someone a level of satisfaction or fulfilment that is free from potential *dukkhā*. In this way, all conditioned objects are deemed in Buddhism to be unsatisfactory as a source of *dukkhā*-free happiness. To see why this is so, we must turn to the Second Noble Truth.

The Second Noble Truth states that the origin of *dukkhā* is *taṇhā*. Broadly, *taṇhā* is an attachment to things being one way rather than another. Specifically, *taṇhā* is a mental disposition whereby one emotionally invests in the satisfaction of a desire, such that there is positive emotion (e.g., joy, relief, excitement) if the desire is fulfilled, or negative emotion (e.g., anger, disappointment, sadness, anxiety) if the desire is frustrated. We first discuss *taṇhā* in its relation to mental *dukkhā*, before turning to *taṇhā* in relation to physical *dukkhā*.

In its relation to mental *dukkhā*, it can be noted that the presence of *taṇhā* is linked to such *dukkhā* in a strong dispositional sense, such that whenever there is mental *dukkhā* (viz., negative emotions) then the disposition of *taṇhā* is

present and whenever *taṇhā* is present, there is the disposition to suffer mentally (in virtue of a desire that can be frustrated). As a mental disposition, the presense of *taṇhā* may not be introspectively obvious to the person who harbours it, which is why the term 'craving' (implying an strong occurrent desire) has inappropriate connotations. For instance, someone who has *taṇhā* towards a life of material prosperity and is getting mostly what he desires – a nice house, good health, car, job and partner – may not overtly feel the tug of his emotional investment in these things; indeed, it does not make much sense to say that he *desires* them while he has them. Yet if any of these items were to be lost or threatened, resulting in his mental *dukkhā* (through negative emotion), his attachment to them would suddenly become obvious. He would not necessarily become any *more* attached to the items than he was before – it is just that the changing circumstances would reveal his existing attachments or *taṇhā* towards them, showing that he had, all along, let his happiness depend on them. The Second Noble Truth asserts that if mental *dukkhā* such as this is to arise, there *has* to be the presence of *taṇhā*. If the man was not in the first place attached to his house, partner, job, etc., then he would not, according to Buddhism, feel any negative emotion at their loss.

Buddhism mentions three classes of item or situation towards which one may harbour *taṇhā*: 'objects of sense-desire', 'becoming' and 'non-becoming'. To have *taṇhā* for objects of sense-desire is to be attached (often implicitly) to the *khandhā* of pleasant feeling that may arise from objects associated with the five sense organs or the mind. It involves emotional investment in such things as food (driven by the pleasant taste sensations), a satisfying career (driven by the pleasant sensations associated with mental fulfilment) and a relationship (driven by alluring emotional sensations of romance). *Taṇhā* with respect to pleasant sensory objects can lead to *dukkhā* if the desired object is threatened, denied or withdrawn. The reason that objects of sense-desire are sometimes classed as *being 'dukkhā'*, even though they feel pleasant, is that their conditioned co-dependence assures them of a lack of immunity from change and decay. To emotionally invest one's happiness in these impermanent objects (e.g., a well-functioning computer) is to hence guarantee the possibility of *dukkhā*, for chances are that either the object of pleasant sensation will not be acquired (leading perhaps to anger or disappointment), it will change or disappear (leading perhaps to grief) or the pleasant sensations will no longer arise in connection with the associated objects: one may get bored, for instance. No conditioned thing, when an object of *taṇhā*, is exempt from the possibility of eliciting *dukkhā* through change, no matter how pleasant it may seem at a particular time. This renders them unsatisfactory as a source of happiness that is completely free of potential *dukkhā*.

To have *taṇhā* for 'becoming' can mean, in Buddhist parlance, that one emotionally invests in being a particular kind of person, for example, as a person who is successful and accepted by one's peers. One wishes to be this kind of person indefinitely, not to go out of existence by means of changing

conditions including eventual death. *Taṇhā* in this context feeds on praise, fortune, love and fame, nurturing what is colloquially known as an ego. *Taṇhā* for existence is not exempt from the possibility of causing *dukkhā*, since the changing world will guarantee that no one, no matter how bright or popular, is immune from rejection, disappointment, embarrassment and ultimately death. Doer-*dukkhā* is often motivated by *taṇhā* for existence (e.g., the desire to be perceived as a success).

On the Buddhist picture, *taṇhā* for 'non-becoming' can arise in connection with being rejected or making a fool of oneself: one yearns, on some level, to not exist. When a person forgets her lines in a well-attended play, wishing that the earth would open up and swallow her, there is temporary *taṇhā* for non-existence – obviously implying *dukkhā*. In the most extreme cases, this can manifest as an urge for suicide, while less dramatic examples can show through in patterns of self-criticism (e.g., 'I am not worthy of any praise'). A more general *taṇhā* for non-existence can pertain not only to one's own person, but to the *khandhā* of unpleasant feeling associated with sensory or mental objects. Just as one can emotionally invest in the presence of pleasant sensations, one may emotionally invest in the absence of unpleasant sensations. To do the latter is to risk suffering the *dukkhā* of wishing an existing situation were otherwise, or that of hoping that an unpleasant situation will not arise. Again, Buddhism claims that while certain objects (e.g., the event of illness) are prone to elicit *dukkhā*, there must also be a vested emotional interest in their non-becoming for that *dukkhā* to have a *mental* dimension. Hence *taṇhā* is required if mental *dukkhā* is to occur. If one did not care at all about the presence of any situation, including the pain of illness, but were perfectly at ease with whatever happened, without *taṇhā* for its non-becoming, then there could not be that layer of negative emotion which arises with wishing that a situation were otherwise.

The Buddhist claim, then, is that *taṇhā* is necessary for the arising of *dukkhā*. I have so far discussed how *taṇhā* (according to Buddhism) could be at the root of the *mental dukkhā* that is behind all unpleasant emotions. But on the Buddhist account, *taṇhā* is also implicated in *physical dukkhā*. Suppose I bang my ankle on the furniture, causing a stab of physical pain. While I might be blamed for being careless, it is not obvious where emotional investment comes into the story. What can it mean for *taṇhā* to be deemed necessary for the pain of physical injury, disease, old age, sickness and death, once we factor out the mental suffering? It is here that we must refer to the broader Buddhist framework of *kamma*, rebirth and *saṃsārā*. Since the idea here is to obtain only a general picture (needed for understanding the Third Noble Truth), I will not elaborate on these concepts further than is necessary.

Basically, Buddhist cosmology maintains that the moral quality of a person's action (where actions include thoughts, speech and physical behaviour) is determined by the amount of *taṇhā* that helps motivate that action.

So long as *taṇhā* is at all present in a person's mind-set, the person's thoughts and so forth, if motivated by that mind-set, will be *kammically* potent.[15] A *kammically* potent action is one that generates future consequences for the person who initiated it (where a person is analysed in Buddhism as a combination of the five *khandhās*). The less *taṇhā* behind an action, the better the *kammic* consequences ('better' meaning a yield of predictably pleasant mental or physical sensations), the more *taṇhā* behind an action, the worse the consequences ('worse' meaning a yield of predictably unpleasant mental or physical sensations). Even a person who dies cannot escape the *kammic* consequences of his former actions – where actions include thoughts on the deathbed – for these actions will condition the rebirth of another body (which is actually a causally conditioned reconfiguration of *khandhās*).[16] The overall conditions of a person's rebirth – whether favourable or unfavourable – will depend on how much wholesome or unwholesome *kamma* has been accrued. But even in a 'favourable' rebirth (such as affluent circumstance), the very having of a body and mind in a physical world will involve physical *dukkhā* at some time or other, whether it be banging one's ankle on the furniture or suffering the flu. This baseline of physical *dukkhā* is thus largely due to the *kamma* of simply *having* a physical body, *kamma* that results from the *taṇhā*-tainted actions of a previous lifetime. And so long as one dies with a mind-set harbouring *taṇhā*, as most of us will according to Buddhism, physical *dukkhā* will be sure to arise in the next birth. This raises the question: does Buddhism believe it possible to escape from all *dukkhā* and indeed, from *saṃsārā*, the potentially endless cycle of birth, death and *dukkhā*? This pertains to the Third Noble Truth.

Before moving onto this, I will note that it is *mental* (not physical) *dukkhā* in its relation to *taṇhā* that will be of relevance to this project. Aside from exegetical purposes, none of my arguments will depend upon the adoption of controversial Buddhist cosmology such as *kamma* and rebirth. The particular interest to be pursued with reference to the Second Noble Truth pertains to the implied relation that Buddhism posits between *taṇhā* and a sense of self – as they relate to the teachings on *anattā* (no-self). For it will be seen in Chapter 3 that Buddhism regards *taṇhā* to not only be the main cause of *dukkhā*, but to co-arise in the mind with a sense of self: a point to be discussed in some detail. A sense of self will hence also, by implication, be at the root of *dukkhā*: a corollary to the Second Noble Truth. The notion of *taṇhā* in relation the *sense of self* will play a major role in later arguments.

The Third Noble Truth states that the cessation of *dukkhā* lies in the cessation of *taṇhā*. The cessation of *dukkhā* involves *nibbāna* (which we shall see is *not* to be construed as out-and-out annihilation). Given that *taṇhā* is at the root of *dukkhā*, the Third Noble Truth functions as a psychological *modus tollens* on the Second Noble Truth: root out *taṇhā*, and *dukkhā* will disappear. In a Western context, this will generally be considered a novel approach to eliminating *dukkhā*.[17] Usually, we seek to avoid suffering by

changing the external conditions that we perceive to be causing the suffering or discomfort. Buddhism does not deny that this can serve as a temporary expedient to stop *dukkhā* – but it is only temporary. So long as our mind-set harbours *taṇhā* – emotional investment in things being a particular way – then sooner or later the conditioned cards will fall in a way that is out of line with our preference, resulting, once again, in *dukkhā*. On the Buddhist picture, the only way to eliminate *dukkhā* once and for all, so that it stops arising altogether, is to eliminate *taṇhā*, the mind-set needed for *dukkhā* to arise.

The elimination of *taṇhā* and hence *dukkhā* is considered by Buddhism to be a real psychological possibility, forming the ultimate goal of Buddhist practice. The term used to designate this goal – the cessation of *taṇhā* and *dukkhā* – is '*nibbāna*' – a term, that we shall see, has more than one nuance.[18] A person who attains *nibbāna* is termed an *Arahant*.

Now, Buddhists distinguish between *nibbāna* while the *Arahant* is living, and *nibbāna* when the *Arahant* dies, a distinction that bears directly upon the difference between mental and physical *dukkhā*. While the *Arahant* is alive, the elimination of *taṇhā* only involves the cessation of mental *dukkhā*. There is still physical *dukkhā* for the living *Arahant*, partly due to the *kamma* of past *taṇhā* which earned him or her a physical rebirth. Lacking *taṇhā*, and hence any emotional aversion to the physical *dukkhā*, the *Arahant* cannot harbour any mental *dukkhā* – that borne of wishing the situation were otherwise. She may thus feel the most excruciating physical pain, but will harbour no mental aversion to the pain; she will not *wish* that the pain would cease. She would become no happier (or less happy) if the pain were to cease. This is what it means to harbour *no* attachment to or *taṇhā* in any states of affairs; such a mode of conscious existence would, by normal psychological standards, be quite extraordinary. Now an *Arahant*, lacking *taṇhā*, will generate no more *kamma* and hence, no conditions for future rebirth into the cycle of *saṃsārā*. So when the *Arahant* dies – an event referred to as *parinibbānā* – there will be no continued and conditioned flux of *khandhās* into another psycho-physical existence. No more perspective will be generated for the witnessing (should there be witnessing):[19] there will be no more subject in relation to objects. With no rebirth, there can be no *dukkhā*, mental or physical. In this context, we can appreciate the full meaning of the Third Noble Truth: the complete cessation of *dukkhā*, mental and physical, must require the cessation of *taṇhā*.

I have mentioned that the term '*nibbāna*' has more than one nuance. In the discussion above, it pertains to the mind where *taṇhā* and (mental and/or physical) *dukkhā* has ceased, or else to the event of *taṇhā* and mental *dukkhā* ceasing. But '*nibbāna*' does not merely denote the cessation of *taṇhā* and *dukkhā*; it denotes something positive as well, since, as I will suggest in Chapter 2, it has something to do with a *mind* that has been liberated from *taṇhā*. The liberated mind of the *Arahant* is, I will argue, the positive thing that

'*nibbāna*' denotes; it is what Buddhism refers to as 'unconditioned' (although the term 'unconditioned' is not a positive but a negative descriptor). So far, I have just talked about how Buddhism regards the conditioned world in terms of the five kinds of *khandhā*. The conditioned and attendable world of such *khandhā*-objects, with their co-dependency, is limited by the aforementioned modes of space, time, quality and relation (although the witnessing element of the 'consciousness' *khandhā* may not be limited in this way). But Buddhist cosmology also alludes to that which is *unconditioned* by these modes of existence and which is not an (attendable) object of witnessing. The mind of the *Arahant* is deemed unconditioned and unattendable in this manner. Given that the *modus operandi* of a subject, witnessing, also seems to elude attention, we may wonder if witnessing *per se* can be surmised to be relevant to the *Arahant's* liberated mind. In Chapters 2 and 3, Buddhist suttas are consulted which imply this to be the case. Moreover, we look at evidence from suttas that suggests there is *something it is like* for the *Arahant* to experience a liberated and unconditioned mind; it is not as if the goal of Buddhist practice is to become a mindless zombie. The implication will be that the witnessing dimension of *nibbāna* – that I will term 'nibbānic consciousness' – has a phenomenal character that is not derived from any object. It is hinted in the suttas that the 'something it is like' tone of nibbānic consciousness involves immeasurable peace and happiness, completely untainted by the presence or possibility of mental *dukkhā*.

The relevance of the Third Noble Truth to this project's enquiry – insofar as it seeks to infer the Buddhist position on self and consciousness – lies in the relation between nibbānic consciousness, ordinary witness-consciousness and sense of self (as construed on my interpretation of Buddhism). I will suggest in Chapters 2 and 3 that one *cannot* get a proper picture of how Buddhism regards the ontology and indeed description of (no)self – and the sense of self – unless one also gets a picture of what is meant by *nibbāna*, in view of how nibbānic consciousness could relate to everyday witness-consciousness (as an element of the fifth *khandhā*). For if it is acknowledged that there is a witnessing dimension to nibbānic consciousness that is unconditioned by time, it will immediately follow that the witnessing dimension to nibbānic consciousness cannot be something that *comes into existence* when a person becomes an *Arahant*. The witnessing must, somehow, be present – perhaps *ever*-present – in a way that is unaffected by time. This invites a number of questions: could the timeless witnessing of nibbānic consciousness be somehow contributing – along with *taṇhā* – to the ordinary sense of self prior to the attainment of *nibbāna qua* event? Could it in fact *be* the very witness-consciousness that is perhaps always present in ordinary conscious states (via the fifth *khandhā*)? Could this ever-present witness-consciousness be instrumental to the mind's liberation from the co-arising *taṇhā* and sense of self – as required for the event of *nibbāna*? I shall present the Buddhist position on consciousness and no-self as answering a resounding 'yes' to

these questions. Such an answer will set the stage for the model of the two-tiered self-illusion to be defended in the book (although the contributing witness-consciousness will not be treated, on our model, as nibbānic).

The Fourth Noble Truth states that there is a path to the cessation of *dukkhā* (and hence to *nibbāna*): the Noble Eightfold Path, which is broadly divided into the threefold practice of 'insight-wisdom' (*paññā*), 'meditation' (*samādhi*) and 'virtue' (*silā*). Being a practice-oriented tradition, Buddhism does not make lofty reference to a mode of flawlessly happy consciousness without prescribing a system of training that will purportedly effect this transformation. Briefly, the practice of 'virtue' seeks to cultivate the kinds of action that are increasingly less motivated by *taṇhā* and more motivated by such qualities as generosity, compassion and equanimity – actions that are 'kammically wholesome', in other words. 'Meditation' pertains to the practice of attentional training whereby the mind is developed, through methods of concentration, to be alert to the arising of *taṇhā*-induced thoughts. In this way, *taṇhā*-induced thoughts and actions can be curtailed before developing into 'unwholesome *kamma*' that will impede progress to *nibbāna*. 'Insight-wisdom' is not book knowledge but pertains, rather, to an immediate and intuitive understanding of the nature of conditioned existence as impermanent, conducive to *dukkhā* and without a self. Buddhism holds that unless one understands conditioned existence in this deep and intuitive way, such that one is no longer compelled to seek happiness from it, one will never attain *nibbāna*, thereby escaping the cycle of rebirth, death and *dukkhā*.

Now, it is not this book's purpose to enquire into the threefold method of training by which the transformation can allegedly be effected. It is the purpose, however, to argue for a Buddhist-derived account of consciousness and no-self (*anattā*) that would *set the stage* for such enquiry. I hope to establish that the self-illusion is 'put together' in such a way that could allow for feasible enquiry into its possible dismantlement via processes outlined in the Fourth Noble Truth.

For now, we have enough background against which to venture forth and interpret the Buddhist position on consciousness and no-self. To this task we now turn.

2
Nibbāna

Introduction

The Third Noble Truth alludes to the ultimate goal of Buddhist practice, *nibbāna*. In a general introduction to the Saṃyuttā Nikāya (a selection of discourses in the Sutta Pitaka belonging to the Pali Canon), Bhikkhu Bodhi writes:

> What exactly is to be made of the various explanations of Nibbāna given in the Nikāyas has been a subject of debate since the early days of Buddhism, with the ground divided between those who regard it as the mere extinction of defilements and cessation of existence and those who understand it as a transcendental (*lokuttarā*) ontological reality. In SN some suttas explain Nibbāna as the destruction of lust, hatred, and delusion, which emphasizes the experiential psychological dimension; elsewhere it is called the unconditioned, which seems to place the stress on ontological transcendence.
>
> (2000, 50)

As conveyed in Chapter 1, I believe that the term '*nibbāna*' can, in the suttas, imply *either* an event of cessation (of *taṇhā* and ignorance) *or* the unconditioned reality, depending upon the context in which the term is used. I also hold that the suttas support a plausible reading in which *nibbāna*, when construed as the unconditioned, is *essentially* experiential, although its full experience is literally beyond imagination. *Nibbāna*, nevertheless, does not *transcend* experience, nor does it fail to exert an influence on daily conscious life. The goal of this chapter is to develop this reading by saying more about what could be meant by *nibbāna* as 'unconditioned' and by inferring the likely relation (based on the suttas) between *nibbāna*, the mind of the *Arahant* and the mind of the ordinary person. In the following chapter (Chapter 3), I infer the contribution that *nibbāna*, via witness-consciousness, could make to the sense of self (within a

Buddhist framework). While my analysis is not meant to be exhaustive, it will provide an angle that is not explicit in the work of many scholars. It will serve as the 'Buddhist reading' that informs this entire project, grounding the 'two-tiered' model of the self-illusion. In Chapter 1, I mentioned that the standard reading of consciousness and no-self in Buddhism tends to overlook the role that unconditioned *nibbāna* could play when construing these notions. My interpretation, by contrast, will take seriously the inferred effect of unconditioned *nibbāna* on the ordinary mind, resulting in some departure from the standard reading.

The primary textual authority for my interpretation of *nibbāna* will be extracts from the Sutta Pitaka. To neutralise any bias in favour of my reading, I rely *mainly* on the translations of Venerable Nyanaponika Therā and Bhikkhu Bodhi, whose commentaries show evidence of *opposing* my interpretation.[1] (The exception will be when I make explicit appeal to a point raised by Thanissaro Bhikkhu). These suttas will sometimes be abridged; I cite only those parts of each sutta that are relevant to the enquiry. Aside from engaging with a brief commentary that represents a mainstream objection to my reading (by Venerable Nyanaponika Therā and Bhikkhu Bodhi), my interpretation as it stands will not rely heavily on secondary sources; I am interested in extracting a picture of *nibbāna* that is based on the suttas themselves. Nevertheless, it should be acknowledged that I am particularly sympathetic to the approaches of Thanissaro Bhikkhu (1993a) and Peter Harvey (1995) whose work I appeal to in places. There are some significant parallels, for instance, between Harvey's and my analysis of *nibbāna*; we both analyse *nibbāna* as involving a kind of unconditioned witness-consciousness or, as he translates it, *discernment* – related to *viññāṇa*, the consciousness *khandhā*. While I approach the concept of *nibbāna* from a more philosophical angle, and with some differences in detail (e.g., in the relation of nibbānic consciousness to ordinary conscious life), my analysis is broadly in harmony with the work of these scholars.

It would be remiss to plough into the analysis without first saying something about the Buddha's own reluctance to dwell on such matters. It was mentioned earlier that the Buddha's primary orientation for teaching was not philosophical but practical: he taught people how to reach the end of suffering. Speculation on the metaphysical nature of *nibbāna* was, in this context, actively discouraged. One reason for this was that being unconditioned, *nibbāna* was deemed ungraspable in any terms that delimited those conditioned objects belonging to the six sense categories, objects that could be attentively perceived by way of eyes, ears, tongue, nose, bodily sensors and mind. Most people (not being *Arahants*) are familiar with only this sensory realm, so to try and imagine the full experience of *nibbāna* would be to superimpose it with qualities that do not apply to it, both confusing and defeating the goal of Buddhist practice. When asked if anything lay beyond the six contact media, Sāriputta, an *Arahant*

disciple of the Buddha, refused to elaborate, saying that it would 'complicate non-complication':

> [Sāriputta]: 'The statement, 'With the remainderless stopping & fading of the six contact-media [vision, hearing, smell, taste, touch, & intellection] is it the case that there is anything else?' complicates non-complication. The statement, '… is it the case that there is not anything else … is it the case that there both is & is not anything else … is it the case that there neither is nor is not anything else?' complicates non-complication. However far the six contact-media go, that is how far complication goes. However far complication goes, that is how far the six contact media go. With the remainderless fading & stopping of the six contact-media, there comes to be the stopping, the allaying of complication.
>
> (AN IV.174)[2]

So are we complicating non-complication by attempting the kind of analysis that is proposed in this chapter? To some degree it will engender proliferation that would be discouraged by those on the Buddhist path. Yet there are mitigating factors. First, the goal here *is* overtly philosophical; I am not purporting to expound Buddhism in any 'spiritual' sense. Second, the tendency for scholars to tiptoe around the concept of *nibbāna* (for fear of saying things they ought not say) has perhaps led to that curious lacuna in standard readings of consciousness and no-self whereby reference to the unconditioned gets left out altogether. Addressing the gap necessitates being more explicit about *nibbāna* in its possible relation to these central Buddhist concepts. Third, what I do say about *nibbāna* will be based, as much as possible, on passages in the suttas. For reasons just mentioned, such suttas are not prolific: yet what is there is both informative and overlooked. While this exercise will involve some joining of dots (I do not pretend that the Buddha explicitly endorsed the view I propose), I will try to keep inference and hypothesis to a minimum. Fourth, just because concepts are used to talk about *nibbāna* (often by saying what it *is not*), it does not follow that *nibbāna* is reduced or made complicated or rendered imaginable by those concepts any more than (to use a famous Zen metaphor) a finger pointing at the moon reduces the moon to the finger. Fifth, some of what is said about *nibbāna* in this analysis may further *clarify* why the Buddha was reluctant to expound upon its nature.

In this chapter, I seek to establish that the suttas strongly support an interpretation whereby the intrinsic nature of *nibbāna* is identical to both the nature of the *Arahant's* percipient mind and to ordinary witnessing that is an element of the consciousness *khandhā* (*viññāṇa*). This will involve two lines of argument. In the first, I seek to establish that (according to Buddhism) the nature of unconditioned *nibbāna* is the nature of the *Arahant's* 'luminous' mind, such that what is true of one is true of the other. To get there, I reason that: (1) unconditioned *nibāna* is real, (2)

unconditioned *nibbāna* is directly experienced by the mind of the *Arahant*, (3) the mind of the *Arahant* is (intrinsically) 'luminous', involving percipience and witnessing, (4) the intrinsic percipient mind of the *Arahant* is identical to unconditioned *nibbāna* (I term the identity 'nibbānic consciousness'), and therefore (5) by indiscernability of identicals, what is true of unconditioned *nibbāna* is true of the *Arahant's* percipient mind and vice versa. (I say more in this section about how to spell out the notion of 'unconditioned'). The second line of argument, which begins in (3) and continues through to (6), seeks to establish that (according to Buddhism) the nibbānic mind – unconditioned witness-consciousness – is intrinsically identical to the witnessing that comes through (even if muted) in the ordinary mind as an element of *viññāṇa*, the 'consciousness' *khandhā*. In establishing this, I attempt to dispel a natural conflict that will arise between the asserted designation of the mind (and consciousness) as unalterable, and those suttas that depict the mind (and consciousness) as being highly changeable. Contrary to popular view, I suggest that the change is extrinsic or relational – a shuffling in the overlay of impermanent objects on the unchangeable mind. On this reading, the path to *nibbāna* will be more accurately viewed as an *uncovering* of the intrinsic mind rather than a literal *development* of the mind. In Section 6, the discussion is brought to bear upon the question of how the intrinsic nibbānic mind could relate to both the mind of the *Arahant* (when viewing objects of awareness) and that of the ordinary person (with *taṇhā* as well as awareness-objects). It is suggested that it is most parsimonious to regard the timelessly ever-present nibbānic consciousness as identical to the witnessing element of the consciousness *khandhā*.

1. Unconditioned nibbāna is real

This premise relates directly to the Third Noble Truth: that there is an escape from *dukkhā* through the cessation of *taṇhā*. Such escape involves reference to the unconditioned. There are several suttas that allude to the unconditioned, including the following, attributed to the Buddha:

> 'There is, bhikkhus, a not-born, a not-brought-to-being, a not-made, a not-conditioned. If, bhikkhus, there were no not-born, not-brought-to-being, not-made, not-conditioned, no escape would be discerned from what is born, brought-to-being, made, conditioned. But since there is a not-born, a not-brought-to-being, a not-made, a not-conditioned, therefore an escape is discerned from what is born, brought-to-being, made, conditioned.'
>
> The born, come-to-be, produced,
> The made, the conditioned, the transient,
> Conjoined with decay and death,

A nest of disease, perishable,
Sprung from nutriment and craving's cord —
That is not fit to take delight in.
The escape from that, the peaceful,
Beyond reasoning, everlasting,
The not-born, the unproduced,
The sorrowless state that is void of stain,
The cessation of states linked to suffering,
The stilling of the conditioned — bliss

(Iti II.16; Iti 37)[3]

2. Unconditioned nibbāna is experienced directly by the mind of the Arahant

This premise implies that *nibbāna* is not a logical abstraction that is merely *inferred* by those who have eliminated *taṇhā* and ignorance from their mind-set; it is something that is 'personally experienced by the wise' (AN III. 55),[4] known directly by the *Arahant*. This also finds support in the previous sutta, for example, 'an escape is *discerned* from what is born, brought-to-being, made, conditioned' (my italics). It is the *mind* of the *Arahant* that ultimately discerns an 'escape' from the bindings of conditioned existence, including those increasingly subtle objects of concentrated meditation states (*jhānās*) listed in the sutta below, and is percipient of *nibbāna*, the unconditioned:

> Once the Venerable Ānanda approached the Blessed One and asked:
> 'Can it be, Lord, that a monk attains to such concentration of mind that in earth he is not percipient of earth, nor in water is he percipient of water, nor in fire ... air ... the base of infinity of space ... the base of infinity of consciousness ... the base of nothingness ... the base of neither perception nor non-perception is he percipient of all these – but yet he is percipient?'
> 'Yes, Ānanda, there can be such a concentration of mind that in earth the monk is not percipient of earth ... nor is he percipient of this world or a world beyond – but yet he is percipient.'
> 'But how, Lord, can a monk attain to such concentration of mind?'
> 'Here, Ānanda, the monk is percipient thus: 'This is the peaceful, this is the sublime, namely, the stilling of formations, the relinquishment of all acquisitions, the destruction of craving, dispassion, cessation, *Nibbāna*.' It is in this way, Ānanda, that a monk may attain to such a concentration of mind.'
>
> (AN X. 6)[5]

The notion of percipience seems close to that of witnessing. It is suggestive of a knowing or observing quality to the mind, even if there is no personal

'knower' or 'observer' as such: an element of conscious apprehension, even if there are (as will be suggested) no *objects* to apprehend.

3. The mind of the Arahant is (intrinsically) 'luminous', involving percipience and witnessing

> This mind, O monks, is luminous, but it is defiled by adventitious defilements. The uninstructed worldling does not understand this as it really is; therefore for him there is no mental development.
> This mind, O monks, is luminous, and it is freed from adventitious defilements. The instructed noble disciple understands this as it really is; therefore for him there is mental development.
>
> (AN I. 49–52)[6]

This sutta suggests that the mind freed from defilements – the mind of the *Arahant* – is naturally or intrinsically 'luminous'. It also suggests, importantly, that all minds (not just that of the *Arahant*) have the inherent nature of luminosity, the defilements being merely adventitious (meaning 'extrinsic, not inherent'). The defilements, rooted in *taṇhā* and ignorance, somehow serve to obscure or dim the mind's natural luminosity. The metaphor suggested by this reading, which I will pursue later in the chapter, is that of the non-enlightened mind being 'covered' by layerings of the defilements, etc., to be uncovered, in its fully luminous state, upon the realisation of *nibbāna*. Mental development will therefore involve a progressive *uncovering* rather than *creating* of intrinsic luminosity. What is the term 'luminosity' meant to convey? Based on the sutta that supported (2), we can surmise that it will involve, at the very least, percipience or witnessing. We can also note that if the intrinsic nature of mind involves witnessing, then an immediate connection is made between *nibbāna* and the everyday mind; the same mind that is percipient of *nibbāna* can be percipient through the defilements. This will prove central to later analysis. The notion of 'luminous' is moreover suggestive of an intrinsic something-it-is-likeness to the nature of mind (as opposed to comatose nothingness), even if its luminous character cannot be attended to as an object and would, in the absence of any objects witnessed, be quite unimaginable.

4. The intrinsic mind of the Arahant is identical to unconditioned nibbāna

Aside from manners of speaking, there is no real canonical evidence to suggest that the mind of the *Arahant* and *nibbāna* are two *separate* things. But there is evidence to suggest that the mind experiencing *nibbāna*, and *nibbāna* experienced, are identical. For instance, in those suttas where the Buddha refers to himself as the 'Tathāgata' meaning 'thus gone', it is the

mind of the Tathāgata that is depicted as freed from any dependency on conditioned things – and freedom from such dependency is the hallmark of *nibbāna*. As Thanissaro Bhikkhu writes (with the inclusion of a sutta of which a portion is cited):

> ... the mind that has attained the goal cannot be known or described from the outside because it is completely free of any dependency — any support or object inside it — by which it might be known. This point forms the context for the dialogue in which the Brahman Upasīvā asks the Buddha about the person who attains the goal.
>
> (1993a, 27)

Upasīvā:
If he stays there, O All-around Eye
unaffected for many years,
  right there
would he be cooled & released?
Would [his] consciousness be like that?

The Buddha:
As a flame overthrown by the force of the wind
  goes to an end that cannot be classified,
so the sage freed from naming (mental) activity
  goes to an end that cannot be classified ...
(S V.6)

According to Thanissaro Bhikkhu (1993a, 2), Buddhaghosa offers the best attempt to describe the etymology of the (Sanskrit) word '*nibbāna*' (in *The Path of Purification*). He construes it as 'unbinding', 'nir'-(un) 'vanā'-(binding) – which Thanissaro Bhikkhu interprets as the freeing of mind from fetters of conditioned existence. In support of this interpretation are many suttas that explicitly refer to the liberation of mind, for example, 'the Tathāgata, by the destruction of the taints, in this very life enters and dwells in the taintless liberation of mind, liberation by wisdom, having realised it for himself by direct knowledge' (AN X. 21).[7]

One must be cautious about taking too literally such dualistic metaphors as the *Arahant* (or the *Arahant's* mind) 'entering' or 'dwelling' in *nibbāna* (as if *nibbāna* were an eternal heaven where the mind or *Arahant* resides) or the mind 'dissolving into' *nibbāna* (as if *nibbāna* were an oceanic consciousness into which the mind dissolves). What dissolves is not the mind but the fetters of *taṇhā* and ignorance that bind it. The event of *nibbāna* is the unbinding of mind: the ontology of *nibbāna*, the mind unbound. In this picture, which has canonical support, *nibbāna* and the unbound percipient mind amount to the same thing.

It is also *philosophically* prudent, within the framework of Buddhism, to identify the mind of the *Arahant* with *nibbāna*. If direct discernment or knowledge of *nibbāna* is to be given the ultimate epistemological status that it purportedly has, then there can be no possibility for error in such discernment. The possibility of error creeps in when there is a metaphysical gap between the appearance of X (to a perceiver) and the reality of X (that is perceived) since the appearance of X may involve a misperception of X's reality. So if the mind is fully percipient of *nibbāna*, the only way to ensure its immunity from misperception is to identify the appearance (the mind's experience) of *nibbāna* with the reality of *nibbāna*. *Nibbāna* – and the experience of *nibbāna* – will be none other than the mind in its pure mode of percipience or witnessing. To construe *nibbāna* as an object – something the mind can pay attention to in principle – would be to separate *nibbāna* perceived from the mind that perceives it, opening up an appearance/reality gap and with it, the possibility for error. So *nibbāna* must be embedded in the principle of percipience itself, with the *full* percipience of *nibbāna* implying the witnessing mind to be completely free from objects of awareness (whether attended to or not). This freedom of the mind from objects would occur in that state of advanced meditative concentration alluded to in AN X. 6 (beyond the *jhānās*) or when the *Arahant's* mind is permanently released upon physical death (in *parinibbānā*).

Concern may arise that the (apparent) witness-consciousness, as I have talked about it in Chapter 1, would be somehow too mundane to ground any unconditioned reality such as that alluded to in Buddhism. I would argue, however, that the concept of witnessing (the *modus operandi* of a subject) shows *prima facie* evidence of being well-suited to this role; it depicts something that is at once both mundane (through its sheer ubiquity) and mysterious (through its elusiveness to attention). For I noted in Chapter 1 that (apparent) witnessing, although cognisant of objects, does not itself come across as being object-like. If the witnessing capacity of a subject could obviously attend to its own witnessing (rendering it an object), then that would presumably make witnessing into something with discernable limitations (such as another subtle object of *jhānā*) – something that could not fit the bill of 'unconditioned'. As that which can attend to sensory-mental limitation, witnessing does not in *itself* seem limited by any observable object qualities. A further reason for supposing that witnessing might (in line with Buddhism) turn out to be unconditioned is that witnessing does not logically require the co-presence of (delimiting) objects that are witnessed. While most of us will be unable to imagine what objectless witnessing is like, it is at least possible to *conceive* of it without contradiction.

Further reflection along these lines may shed light on why it could be altogether *impossible* to imagine the full percipience of *nibbāna* – and why, hence, the Buddha was reluctant to talk about it. Normal conscious states, insofar as they contain objects, seem polarised into a field of attention and inattention, as well as subject and object. Objectless 'nibbānic' witnessing

would presumably lack any such polarisation and so would not resemble the familiar structure of consciousness. The act of trying to imagine anything (in terms of what it is like) involves the framing of an imagined item as an *object* – something that is attended to in the 'mind's eye' of a *subject* – but the very structure of consciousness presupposed by this act would negate what is essential to the experience of objectless and subjectless *nibbāna*. Hence trying to imagine the full experience of *nibbāna* would be even more futile than a colour-blind person trying to imagine what colour experience is like; at least colours resemble other sense experiences insofar as they are objects that can be focused upon by a subject.

Once it is acknowledged that the experience of *nibbāna* is beyond imagination, the use of metaphor may be of aid in gaining a better abstract understanding. Illumination metaphors are often used to illustrate witness-consciousness, so one may be tempted, along these lines, to conceive of 'nibbānic consciousness' (as I shall call it) as akin to a boundless ocean of sparkling light. I think this is too crude. Thanissaro Bhikkhu (1993a, 1–4) points out that the releasing of the *Arahant's* mind (upon bodily death) is often compared, in the suttas, to a flame going out – when, in those times, fire was considered to subsist everywhere in its latent potentiality, too subtle to be perceived. A more suitable analog of this, on the metaphor of light, may be that of light beams: luminous by nature but invisible to us unless hitting the eye, whether directly or by means of reflection from something else such as a cloud of dust. Although invisible, the beams of light are far from nothing; they exist as a potential *source* of visible illumination (and importantly nothing about the light *intrinsically* alters with the addition or subtraction of dust). Similarly, I am inclined to think that 'nibbānic' witness-consciousness, although 'luminous' by nature (percipient), does not present as an active 'beam' of awareness (in seeming to emanate from the spatio-temporal perspective of a subject) unless objects, witnessed through the six senses, are there to 'reflect' and 'direct' it. Then there is the appearance of a familiar structured field of consciousness (although nothing about the witness-consciousness *intrinsically* alters with the addition or subtraction of objects).

The problem with the 'ocean of sparkling consciousness' metaphor is that it seems to surreptitiously blanket the universe with a structured field of consciousness, as if objects were scattered everywhere to reflect a subject that is everywhere. But while there are no objects to 'reflect' pure nibbānic consciousness, on the metaphor that I suggest, it is not to be construed as an ocean of nothingness either. It is the ultimate source of 'bright' conscious experience that extrinsically differs from the channelled and focused consciousness-with-objects. The overt manifestations of object-directed consciousness, I suggest, are what Buddhism alludes to as the 'consciousness' *khandhā*. When objects of awareness disappear (such as upon physical death of the *Arahant*), the *khandhā* no longer exists *as such*. Yet its intrinsic tie to nibbānic consciousness – the element of witnessing – does not cease to be.

In the final section of this chapter, I expand further on the relation that nibbānic consciousness could bear to witness-consciousness (as an element of the consciousness *khandhā*), both in the mind of the *Arahant* (when perceiving objects) and the mind of the person bound by *taṇhā*.

5. By indiscernability of identicals, what is true of unconditioned nibbāna is true of the Arahant's mind and vice versa

So far I have argued that Buddhist suttas (supplemented with philosophical extrapolation) support an identity of unconditioned *nibbāna* with the luminous mind of the *Arahant*, an identity I refer to as 'nibbānic consciousness'. By the indiscernability of identicals, what is true of *nibbāna* must be true of the *Arahant's* mind (and vice versa) and I have already said something about what it could mean for *nibbāna* to be identical to the *Arahant's* percipient mind. As something directly *experienced*, *nibbāna* (in its full percipience) is not conveyed as the mere cessation of craving (ā la Kalupahana, 1976, 69–83), but neither is it likened to an infinite ocean of 'active' consciousness-with-objects; it is not, indeed, like anything imagined. I have not, however, said much about what it could mean for such a mind/*nibbāna* to be *unconditioned*: something that I will now expand upon. This will have particular implications for those commentaries that mistakenly, I believe, designate the luminous mind as intrinsically (rather than extrinsically) changeable.

In Chapter 1, I spoke of the five *khandhās*, in Buddhism, as being conditioned by the ('Kantian') parameters of space, time, quality and relation. In the suttas there are passages to suggest that *nibbāna* is, by contrast, unconditioned by these parameters. The depiction of *nibbāna* (or what we can infer from its depiction) bears many similarities to that of Immanuel Kant's 'noumenal subject' (1787)[8] and in what follows I will draw attention to some similarities and differences between the two. What applies to *nibbāna* will, *ipso facto*, apply to the intrinsic luminous mind of the *Arahant*.

5.1. Unconditioned by quality

In the previous chapter, the term 'quality' was stipulated, in a Buddhist context, to describe anything that can be attended to by means of sensory or mental faculties – which covers all of *conditioned* existence. It hence pertains to all the empirical qualities that characterise or can be known through the five sense modalities (discernible objects with colours, shapes, sounds, tastes, tactile and proprioceptive qualities) as well as qualities and concepts which characterise mental deliverances such as emotions, perceptions, thoughts and mathematical concepts. This realm of conditioned existence, perceivable by means of the six spheres of sense-contact, is

described as 'differentiated'. *Nibbāna*, by contrast, is depicted as the allaying of differentiation. For example, while in AN IV. 174 Sāriputta spoke of that beyond the six sense-contacts as 'non-complication' he says in the previous sutta:

> However far the six spheres of contact go, that is how far differentiation goes. However far differentiation goes, that is how far the six spheres of contact go. With the remainderless fading and stopping of the six spheres of contact, there comes to be the stopping, the allaying of differentiation.
> (AN IV. 173)[9]

Insofar as it is unconditioned by qualities pertaining to the six spheres of contact, *nibbāna* seems notably parallel to Kant's (1787, A404) 'noumenal subject', namely, as 'not a real whole but a simple'. I interpret Kant to mean here that the noumenal subject, as a simple, is undifferentiated by any empirical or conceptual determination. But while Kant believes the deliveries of sense and reason to exhaust the scope of human knowledge and experience (perhaps rendering the noumenal self to be little more than logical abstraction) this does not seem to be true with respect to *nibbāna*. The idea that *nibbāna*, beyond the differentiated sense-realm, can be directly experienced and understood – (and as I have argued, *not* as a separate *object* of experience) – sets it apart from Kant's humanly unknowable noumenal subject:

> Therefore Bhikkhus, that base should be understood, where the eye ceases and perception of forms fades away. That base should be understood, where the ear ceases and perception of sounds fades away ... That base should be understood where the mind ceases and perception of mental phenomena fades away. That base should be understood.
> (SN XXXV. 117)[10]

> 'Through dispassion [towards the five *khandhās* of conditioned existence] [his mind] is liberated. When it is liberated there comes the knowledge: 'It's liberated.'
> (SN XXXV. 28)[11]

The noetic/experiential dimension to *nibbāna* is furthermore hinted at as being extremely positive in hedonic tone, for example, in various epithets of *nibbāna*: 'the sublime, the auspicious, the wonderful, the amazing' (SN IV. 43),[12] in Iti II.16; Iti 37 as 'the stilling of the conditioned, bliss'[13] and in AN IX. 34 as 'happiness' (*sukhā*). Yet for reasons already outlined, one must be cautious not to juxtapose *nibbāna* with an imagined experience of happiness; such an exercise will falsely reify *nibbāna* as an object of consciousness.

There is a further consequence of significance here. Anything pertaining to the 'allaying of differentiation' carries connotations of an *absolute*

unity – absolute in the sense that there is no observable differentiation to suggest a lack of unity. Experienced as it is in itself *sans khandhās*, with no sense of a subject/object (or attention/inattention) division, nibbānic consciousness cannot therefore seem like a disunified or fragmented witnessing but must carry a sense of absolute intrinsic unity. In the following chapter, I consider how nibbānic consciousness – even if extrinsically 'covered' through the presence of spatio-temporal *khandhās* – could also present a unity with respect to these *khandhās*. In later chapters, and outside the context of nibbānic consciousness, I will argue that witness-consciousness brings to ordinary conscious experience a unity that is not only inferred (in the manner of Kant's noumenal subject) but directly experienced.

5.2. Unconditioned by space

For Kant, human 'intuitions' of space and spatial qualities originate not from the mind-independent world (as would naively seem) but from the noumenal subject – that humanly unknowable entity which must be postulated if law-like conscious experience is to be possible. Kant (1787, A 404) maintains that this entity cannot in itself be conditioned by space, but must rather present a unity with respect to it. With respect to *nibbāna*, its being unconditioned by space can be found in a simple implication; if something is unconditioned by quality in virtue of being beyond the 'six spheres of contact', then it cannot be conditioned by space, since spatial limitation bespeaks of discernible qualities that can be known through these spheres. That is, being conditioned by space involves an object's restriction by spatial parameters (such as position and shape), which can be verified through sensory observation. In Buddhism, the *khandhā* of physical form, with its elements of 'earth, fire, water and wind' (pertaining to density, heat, fluidity and airflow) is allocated the 'space-filling' niche. Since this renders the spatial physical *khandhās* describable with reference (at the very least) to those sensory-mental modalities of sight, touch and perception, they must also be conditioned by quality. Lacking any such discernible quality, *nibbāna* cannot therefore be conditioned by space.

5.3. Unconditioned by time

> There are, O monks, three conditioned marks of the conditioned. What three? Its origination is discerned, its vanishing is discerned, its change while persisting is discerned. These are the three conditioned marks of the conditioned.
>
> There are, O monks, three unconditioned marks of the Unconditioned. What three? No origination is discerned, no vanishing is discerned, no change while persisting is discerned. These are the three unconditioned marks of the Unconditioned.
>
> (AN III. 47)[14]

Often referred to in suttas as 'the deathless', *Nibbāna* is clearly portrayed as unconditioned by time. Kant's (1787, A 404) noumenal subject is 'the unconditioned unity in the plurality in time' in that it is 'not numerically different at different times but one and the very same subject'. In context of Kant's metaphysics, this means that the noumenal subject is outside of time, rather than continuing forever in time. Since Kant regards time, like space, to originate from the noumenal subject, he sees the noumenal subject as exempt from temporal plurality; it presents a unity by not being numerically different at different times (and so not perduring). Like Kant's noumenal subject, *nibbāna* is to be understood as being eternal in the sense of being *outside of* time, rather than as existing forever *within time*.

Unlike Kant's noumenal subject, however, the unifying presence of timeless nibbānic consciousness can, as we have surmised, be discerned through direct experience. This invites a potential objection. Suppose that at 4 pm an *Arahant* enters a meditative state where he is fully percipient of *nibbāna* and then at 5 pm emerges from the state. Does this not imply that *nibbāna* is conditioned by time – that it is an impermanent *state* he has gone into and will emerge from? I believe not. It is reasonable to suppose that what makes 'entering into *nibbāna*' *appear* like a temporal state has everything to do with the extrinsic dispersing and re-grouping of awareness-objects (viz., groupings of the *khandhās* within his field of consciousness) that *are* conditioned by time. To use an earlier metaphor, such an event is perhaps akin to the disappearance and reappearance of dust, which manifestly reflects the light's intrinsic luminosity, seeming to affix it to a place and time. The coming and going of the dust has no intrinsic effect on the light although it creates the outward impression of change. On another metaphor, the presence of *nibbāna* may be akin to that of the sun being unaltered by clouds that cover it (the cloud-covering metaphor perhaps more appropriate to the idea of *taṇhā* and ignorance appearing to dim the luminous mind). On either metaphor, the effect of the *khandhās* is not to create conditions that impinge upon the reality of nibbānic consciousness, but to create conditions that may in some way *limit* its full percipience, whether the limitation is through divergence and reflection by awareness-objects or through obscuration by *taṇhā* and ignorance. The *nibbāna* would be conditioned by time if, like a cloud or dust particle, it were actually brought into existence upon that moment, but the suttas clearly imply that this is not the case. *Nibbāna* – indeed, nibbānic consciousness – might be therefore described as being 'timelessly ever-present'. The timeless aspect suggests that nibbānic consciousness, in its undifferentiated (hence non-plural) unity, is intrinsically unaffected by any time-bound configuration of *khandhās*. The ever-presence suggests that at any given *khandhā*-moment (or 'k-moment' for short), the unified *nibbāna* will be present – even if not obviously so.[15] And in the absence of any k-moment (or where there is no passing of time) *nibbāna* will be present.

We should therefore not be misled by the convention of referring to a subject or person as 'coming to experience *nibbāna* at time *t*' anymore than we should be misled by talk of 'the sun rising or setting'. The timeless *nibbāna* he or she 'partakes in' is not, intrinsically, a mere *state* or *activity* of his (*qua* subject's) mind that somehow comes into existence (like a state of excitement) when conditions, or the lack of them, are ripe. It is rather the underlying *nature* of mind that is being unveiled, a mode of unconditioned reality that does not intrinsically depend upon the mental or physical states that delimit the conditioned parameters of a subject or person. Yet the composite impression of luminous mind plus overlying *khandhā*-activity will be that of a mind in flux, sometimes directed towards one object, sometimes another (even more changeable if under the influence of *taṇhā*). Given what has been argued so far, it is very important to construe such alteration of the mind to be *extrinsic* (relational, to do with adventitious movement or 'coverings') rather than *intrinsic*. This is how I would interpret such phrases as 'the mind ceases' or 'the mind changes' (what ceases in pure *nibbāna* is the mind in its *relation* to *khandhās*) as in the following sutta:

> No other thing do I know, O monks, that changes so quickly as the mind. It is not easy to give a simile for how quickly the mind changes.
>
> (AN I. 48)[16]

Commenting on this sutta in relation to that on the luminous mind (already cited), Venerable Nyanaponika Therā and Bhikkhu Bodhi write:

> 'The fact that this expression 'luminous mind' does not signify any 'eternal and pure mind-essence' is evident from the preceding text, [above] in which the mind is said to be extremely fleeting and transitory.'
>
> (1999, 278)

This interpretation of change as being *intrinsic* to the mind typifies the view taken by scholars on the underlying mind in Buddhism (significantly influencing how the concept of no-self comes to be interpreted). When change is interpreted as *extrinsic*, however, a completely different picture, as argued for in this chapter, emerges; the luminous percipient mind *is* eternal, in the sense of being unbroken and ever-present, unconditioned by time. In addition to what has so far been argued (via the indiscernability of identicals) it makes no logical sense, in context of Buddhist teachings, to construe the ultimately discerning mind – the principle of percipience – as itself transitory and fleeting. Buddhism makes much of the idea that the conditioned nature of phenomena, including its impermanence, is to be apprehended and understood through direct experience (not merely inferred or remembered). So let us for argument's sake cast the discerning mind as intrinsically impermanent, as conditioned; this too must be known

through direct experience. Such a mind, in order to directly experience (and hence know) its own impermanence, would have to be percipient of its own fleeting nature. That means that it would have to be present while it directly discerns its own fleeting moments of absence (as well as presence). But then if present to its own absence, it cannot actually be *absent* during those moments; we arrive at a contradiction. Therefore the mind cannot directly discern its own fleetingness, if it is indeed fleeting. So, in keeping with the canonical endorsement of the idea that the mind directly (rather than inferentially) discerns the conditioned nature of phenomena, we have to rescind on the idea that the discerning mind is in itself fleeting and conditioned. The discerning percipient mind cannot, in itself, be transient and fleeting; it cannot be conditioned; it must be unconditioned, and hence (in line with everything we have said), *nibbāna*.

5.4. Unconditioned by relation

To be unconditioned by relation, on Kant's (1787, A 404) account, is to be 'not inherent [in something else] but self-subsistent', which would imply, as with the noumenal subject, an unreliance on anything else for existence. The unreliance of *nibbāna* upon the *khandhās* can be inferred from the fact of its being unconditioned by time. Since *nibbāna* is not time-bound, it cannot depend upon any time-bound entity, viz., the *khandhās*. If it did depend upon the *khandhās*, then it would, like the decaying *khandhās*, be itself subject to decay. But *nibbāna* is portrayed as unequivocally exempt from those laws of conditioned co-dependence (both synchronic and diachronic) that govern the five *khandhās*.[17] Thanissaro Bhikkhu (1993, 28–29) has noted that the complete freedom of the *Arahant's* mind from dependency upon any object renders it indescribable both 'from the outside', and from the perspective of the *Arahant* who experiences it. There are simply no quality-restricted criteria by which it *could* be positively described:

> When all phenomena [viz., the conditioned *khandhās*] are done away with all means of speaking are done away with as well.
>
> (SN V. 6)[18]

Nibbāna hence emerges as every bit unconditioned as Kant's noumenal subject – and yet its identity with the percipient mind of the *Arahant* would lend it, unlike with Kant's noumenal subject, an intrinsically experiential (although literally unimaginable) dimension.

6. Nibbānic consciousness and the khandhās

We now examine in more detail how Buddhism (as so far interpreted) might plausibly relate the intrinsically timeless, objectless *nibbāna* to conscious 'states' where there is the co-presence of *khandhās*. For this purpose, two

important groups of people will be considered: (1) the *Arahant* (in whom *taṇhā* has ceased) and (2) the ordinary person with *taṇhā*. In both groups, it will make the most parsimonious sense to suppose that the *khandhā* of 'consciousness' (*viññāṇa*) is the vehicle through which nibbānic consciousness, via the witnessing dimension, makes itself known, even if 'diverted' or 'dimmed' through co-presence with other *khandhās*. The fact that the consciousness *khandhā* is a designated part of the conditioned world will be due to its object-directedness. It is akin to 'the mind' when construed in its relational (hence changeable) mode; it ceases to exist (in its relational mode) upon the full percipience of (objectless) nibbānic consciousness. This will explain those *suttas* that refer to consciousness as 'ceasing' upon *parinibbānā*.

6.1. The Arahant with 'proximate nibbānic consciousness'

We have surmised that for an *Arahant* to be fully percipient of *nibbāna* there cannot be the co-presence of experienced *khandhās*. Nibbānic consciousness as it is in itself will be subjectless and objectless. Upon the *Arahant's* physical death (*parinibbānā*) the psycho-physical *khandhās* will not reconfigure into another birth. The nibbānic consciousness formerly associated with *khandhās* of the *Arahant* will be fully percipient, free from any confines of space, time, quality and relation. Now during their psycho-physical lifetime, the mind or consciousness of an *Arahant* will come into regular contact with objects of awareness as they engage with the world. As Harvey (1995, 197) points out, this implies that the objectless *nibbāna as* objectless 'is not experienced by the Arahat all the time'. However, Harvey (1995, 192) also notes that 'he or she can repeatedly re-experience it before entering it for a final time at death ... The Arahat's full experience of *nibbāna*, as a state in which the personality-factors [*khandhās*] stop, might be seen as his 'participating in' this timeless reality'.

While the *taṇhā*-free *Arahant* has ready access to timeless, objectless nibbānic consciousness, insofar as their *khandhās* can temporarily 'stop' to allow for its full experience, there are plenty of *k-moments* where *khandhās* are present to their field of consciousness. In these k-moments, where percipience of *nibbāna* is not entire, how would nibbānic consciousness reveal itself? In the earlier-cited metaphors, one might construe the *Arahant's* mind, with its lack of *taṇhā* and ignorance, as akin to a sun uncovered by clouds. So long as objects are perceived, such a mind will retain the impression of occupying a distinct beam-like spatio-temporal perspective. This impression will however be viewed as being merely *extrinsic*; for there will be the central realisation – gained through 'partakings' in objectless consciousness – that witnessing is not *intrinsically* confined to such a perspective (just as light is not confined to the shape that a sparkle of dust may impinge on it).

Psychologically speaking, what comes across from the accounts of reported *Arahants* is that even while interacting in the world, their minds, never far from full percipience of *nibbāna*, are replete with aspects whose source can be

most parsimoniously attributed *to nibbāna*, most obviously via the *khandhā* of (object-oriented) consciousness itself. Having realised the goal of the Noble Eightfold Path – that of eliminating *taṇhā* and experiencing *nibbāna* – they can never revert to their former state of *taṇhā* and of not knowing what *nibbāna* is fully like.[19] (Perhaps in a similar fashion, one who has clearly 'seen through' a magic trick can never be duped by it again). Realising this ultimate Buddhist goal is said to involve the deepest understanding of conditioned phenomena, meaning that the mind, fully realising the impermanence of *khandhās*, is never again drawn to seek happiness from them:

> He reflects on it thus: 'This gain that has come to me is impermanent, bound up with suffering, subject to change'. And so he will reflect when loss and so forth come upon him. He understands all these things as they really are, and they do not engross his mind. Thus he will not be elated by gain or dejected by loss ...
>
> (AN VIII. 6)[20]

The mind of the *Arahant* thus fully comprehends the conditioned *khandhās* as *aniccā*, *dukkhā* and *anattā* – that is, as impermanent, as an unsatisfactory source of *dukkhā*-free happiness, and as not pertaining to a self (more on this in the next chapter). What is termed as 'ignorance' (*avijjā*) about conditioned existence – ignorance that has one continuing to seek *dukkhā*-free happiness from the impermanent *khandhās* (an unfulfillable goal) – is hence said to be lost, resulting in full 'wisdom' (*paññā*) with respect to this factor. The *Arahant* also apprehends the unconditioned nature of *nibbāna* as not *aniccā* and not *dukkhā* (but still *anattā* since there is no place for a self – more on this in Chapter 3). Upon apprehending the unconditioned, there is complete release from *taṇhā*; the *Arahant* is no longer emotionally buffeted by the 'vicissitudes of life'. To continue the previous sutta:

> Loss and gain, disrepute and fame,
> Praise and blame, pleasure and pain –
> These things are transient in human life,
> Inconstant and bound to change.
> The mindful wise one discerns them well,
> Observant of their alterations.
> Pleasant things do not stir his mind
> And those unpleasant do not annoy him.
> All likes and dislikes are dispelled by him,
> Eliminated and abolished.
> Aware now of the stainless, griefless state,
> He fully knows, having gone beyond.
>
> (AN VIII. 6)[21]

In many reports, the mind of a living *Arahant* is not only free from *taṇhā*, mental *dukkhā* and ignorance about conditioned reality, but is radiantly and constantly happy in the knowledge of its real nature as unconditioned (with untainted happiness deemed inherent to the mind's very nature). Such a mind's knowledge of its real unconditioned nature is never forgotten while engaging with objects of the world; lacking *taṇhā*, there is no mental *dukkhā*. The *Arahant's* sense of witness-consciousness is said to be very powerful, present-centred and constant, with awareness never 'lost' in the imagined time zones of thought. Buddhism maintains that being regularly lost in thought feeds a mindset with *taṇhā*, which, it is claimed, is pivotal to maintaining the sense of being a separate self. In the next chapter it will be suggested (on the Buddhist analysis) that the witnessing of a usual *taṇhā*-ridden person, while sourced in nibbānic consciousness, does *not* reflexively know its real nature as unconditioned – because it assumes, reflexively and mistakenly, that it is a separate self. Lacking the co-arising *taṇhā*, *Arahants* are said to be without a sense of self. We can surmise that a pull of *taṇhā* will strongly condition the arising of mental *khandhās* (such as thoughts) in the field of awareness – hardly conducive to ready partakings in objectless consciousness. Freed from *taṇhā*, however, the mind of the *Arahant* will be naturally 'with' whatever is happening in the present – not compulsively looking elsewhere (for example, to past or future scenarios) for gratification. This may considerably free up their mind to 'partake in episodes' of objectless consciousness. Because of its ready proximity to full nibbānic consciousness, I refer to the *Arahant's* witness-consciousness with presence of *khandhās* as 'proximate nibbānic consciousness'.

6.2. 'Pre-nibbānic consciousness' of the ordinary person

While 'proximate nibbānic consciousness' denotes the witness-consciousness of an *Arahant* who is observing objects, I allocate the term 'pre-nibbānic consciousness' to denote the witness-consciousness of the person whose mind harbours *taṇhā* – no matter how much or how little. As before, I surmise that the witness-consciousness, intrinsic to *nibbāna*, shows through the everyday mind as an element of *viññāṇa*, the consciousness *khandhā*. It is just that with the everyday mind, the full percipience of witnessing is inhibited not only by the co-presence of awareness-objects (which structure the field of consciousness into subject/object, attention/inattention) but by the (extrinsic) dimming and distorting of consciousness through coverings of *taṇhā* and ignorance (as clouds cover the sun). To treat the witnessing element of *khandhā*-consciousness as something altogether unrelated to the witnessing of nibbānic consciousness is to multiply entities beyond what is necessary, as well as leave unanswered the question: 'if *nibbāna* involves the principle of witnessing, and is timelessly ever-present, then what effect could it exert in daily conscious life?' It makes the best parsimonious sense, therefore, to construe Buddhism as upholding no *intrinsic* difference between the 'pure'

nibbānic witness-consciousness of an *Arahant* and the witness-consciousness of a criminal. The difference between 'pure', 'proximate' and 'pre'-nibbānic consciousness is only *extrinsic* – to do with the 'covering' or 'reflecting' of *khandhās* and/or *taṇhā*. A criminal has the potential to attain *nibbāna* – a potential to be realised not by his 'watering the seed' of a latent nibbānic consciousness so that it 'blossoms into' a fully fledged version, but by his eradicating *taṇhā* so that conditions become ripe for an 'uncovering' of the timelessly ever-present nibbānic consciousness, intrinsic to his mind. Our enquiry will now turn to the extent to which, on this Buddhist position, the underlying objectless nibbānic consciousness would likely be revealed in a mind with the coverings of *taṇhā*. How much, if at all, would pre-nibbānic consciousness, with the co-presence of *taṇhā*, resemble the *taṇhā*-free objectless nibbānic consciousness? Obviously the resemblance cannot be veridical – otherwise there would be no need for the Four Noble Truths to help one's mind to realise *nibbāna*.

From reflections on this chapter we can note an immediate similarity: *everyday witnessing* – the *modus operandi* of a subject that is an element of the consciousness *khandhā* – will resemble that of nibbānic consciousness. While Buddhism does not make any notion of witnessing explicit – since the focus of practice is on discerning the nature of *khandhā*-objects as *aniccā*, *dukkhā* and *anattā* – witnessing as 'that which knows' objects as *aniccā*, *dukkhā* and *anattā* must be assumed in the very possibility of such practice. For if there is to be any progression in 'wisdom' there must be an aspect to consciousness that can progressively *know* the conditioned nature of *khandhās* and of course, know itself reflexively as the unconditioned nibbānic consciousness.

Resuming our cloud-covered sun analogy, if the sun is analogous to the pure objectless nibbānic consciousness, the witnessing inherent to pre-nibbānic consciousness may be compared to the brightness of clouds from the sun behind them. We may now wonder whether Buddhism would regard any further aspects of nibbānic consciousness to be 'brightening the clouds'. For instance, does Buddhism regard the ordinary conscious state of the average person to harbour any indications – even if somewhat muted or distorted – of a sense of timeless (or unbroken) presence, non-dependence upon objects, unchangingness, non-differentiation or unity, and intrinsic happiness? To determine this, we must now consider what Buddhism has to say about the self which will be the topic of Chapter 3.

3
The Definition and Status of Self in Buddhism

Introduction

Buddhism is famous for its alleged claim that the self does not exist. In this book I seek to defend the Buddhist 'no-self' principle (*anattā*) on Western philosophical grounds. The goal of this chapter, therefore, is to gain a plausible picture of what is involved in the Buddhist 'no-self' principle. While I agree with Thanissaro Bhikkhu (1993b) that *anattā* is presented in the suttas as more a *strategy* (for letting go *dukkhā*) than an ontological assertion, I will be approaching the matter from an ontological perspective. As such, there will be a fair amount of inference. The emerging picture will be one that takes seriously the role of 'pre-nibbānic' witness-consciousness, inferred from Chapter 2.

The first question to be addressed in this enquiry is: 'what is the self whose existence Buddhism denies? How do we define it?' With their pedagogical bent, the suttas do not have an explicit answer. Extracting a clear and plausible definition of self from the suttas will therefore take up a good section of this chapter; I have found no satisfactory definition in secondary sources. Now given that losing the *sense* of self has been correlated with ditching *taṇhā* on the road to *nibbāna*, the self, as construed in Buddhism, must be something we could plausibly *have* a sense of being. It cannot be anything too grandiose. While 'immortalised' notions of self are also mentioned in the suttas (and the existence of such selves implicitly rejected), these concepts of self are peripheral to the main concerns of Buddhism and will not be discussed in this chapter.

Once defined, it will not be hard to infer that Buddhism would deny existence to such a self *as a whole*. It is when analysing various features ascribed to the self (and known to us through the *sense* of self) that more contentious issues arise. Are each of its ascribed features regarded as equally unreal? The importance of the analysis in Chapter 2 now comes to the fore. Several features such as unity and unbroken persistence, which scholars standardly dismiss as lacking reality in Buddhism will be most likely

grounded in nibbānic consciousness via the witnessing element of the consciousness *khandhā*. As such, the Buddhist suttas support a reading which casts them *not* as mere mental fictions (a la Hume's bundle theory), but as having a measure of independent reality – or so I will argue.

1. Defining the self in Buddhism

1.1. Ownership and identification as reciprocal assumptions of self

To understand what 'self' means in Buddhism, we must get some grasp of what is involved in the concepts of *identification* and *ownership* (expressed in terms of 'me' and 'mine') as they are used in a Buddhist context. I now spend some time unpacking them, a task that will run into Chapter 4.

In essence, the concept of identification, as I glean it from Buddhism, conveys a reflexive and implicit assumption, on the part of the subject, that various psycho-physical attributes (combinations of *khandhās*) are in some way assimilated to itself.[1] Ownership can be regarded in Buddhism as reflecting a broad mode of such identification, where the subject assimilates to itself those *khandhās* that underpin the role 'owner of X'. This assumed identity of the subject as 'owner' is evidenced through the assumption of certain other *khandhās* to be 'mine' or 'belonging to me'.

We will see that Buddhism regards the ownership mode of identification to be pivotal to the notion of self, such that if there is a self (or 'me'), then there is what belongs to a self ('mine'), and if there is what belongs to a self ('mine'), then there is a self ('me-as-owner'). From the suttas below, it will be discerned that the aim of Buddhist practice is to rid one's mindset of any *sense* of 'me' and 'mine'. The cognitive categories of 'me' and 'mine' – indicating the self's identity as some kind of separate bounded owner – will turn out, crucially, to ground the self's status as lacking reality. The goal of Buddhist practice will thus be described as one in which the subject becomes directly attuned to the reality of no-self, by no longer making false assumptions of ownership and self-identity. By failing to assume the *khandhās* to be a part of itself, the subject will eliminate the psychological platform of a supposedly separate 'me' from which other *khandhās* are regarded as 'mine'. When the sense of self is eliminated in this manner, then so is rebirth into *saṃsārā*, with its endless round of *dukkhā*.

We should be reminded that the definition of *subject* from Chapter 1 is *witnessing as it presents from a psycho-physical perspective*. It is critical to the following discussion that this notion of subject does *not* involve appeal to ownership or identification and so is not the same as the notion of self depicted in Buddhism – although the Buddhist notion of self will incorporate it. On the Buddhist position, we are to understand that the witnessing subject makes the (deeply mistaken) assumption of *being* a self through its very act of *assuming* various *khandhās* – including those that help lend it a psycho-physical perspective – to be 'me' (hence integrated with its

existence) or 'mine' (hence belonging to it). Let us then consult some of the relevant suttas (whose themes are repeated in other suttas) that support this reading. We should be reminded that only the relevant parts to each sutta are quoted:

> 'Bhikkhus, there being a self, would there be for me what belongs to a self?'
> 'Yes, venerable sir.'
> 'Or, there being what belongs to a self, would there be for me a self?'
> 'Yes, venerable sir.'
>
> (MN 22)[2]

> 'Just as a dog, tied by a leash to a post or stake, keeps running around and circling around that very post or stake; in the same way, an uninstructed, run-of-the-mill person – is not well-versed or disciplined in their Dhammā – assumes [*khandhās* (form, feeling, perception, mental formations, consciousness)] to be the self, or the self as possessing [*khandhās*], or [*khandhās*] as in the self, or the self as in [*khandhās*].
>
> He keeps running around and circling around that very form ... that very feeling ... that very perception ... those very fabrications ... that very consciousness. He is not set loose from form, not set loose from feeling ... from perception ... from fabrications ... not set loose from consciousness. He is not set loose from birth, aging, & death; from sorrows, lamentations, pains, distresses, & despairs. He is not set loose, I tell you, from suffering & stress.
>
> [The mode to aspire is where one] doesn't assume [*khandhās* (form, feeling, perception, mental formations, consciousness)] to be the self, or the self as possessing [*khandhās*], or [*khandhās*] as in the self, or the self as in [*khandhās*].
>
> He doesn't run around or circle around that very form ... that very feeling ... that very perception ... those very fabrications ... that very consciousness. He is set loose from form, set loose from feeling ... from perception ... from fabrications ... set loose from consciousness. He is set loose from birth, aging, & death; from sorrows, lamentations, pains, distresses, & despairs. He is set loose, I tell you, from suffering & stress'.
>
> (SN XXII.99)[3]

> [On the Arahant] 'In the same way, a monk investigates form, however far form may go. He surveys feeling ... perception ... fabrications ... consciousness, however far consciousness may go. As he is investigating form ... feeling ... perception ... fabrications ... consciousness, however far consciousness may go, any thoughts of 'me' or 'mine' or 'I am' do not occur to him'.
>
> (SN XXXV. 205)[4]

> '... How do you construe this, monks: If a person were to gather or burn or do as he likes with the grass, twigs, branches & leaves here in Jeta's Grove, would the thought occur to you, 'It's *us* that this person is gathering, burning, or doing with as he likes'...'
>
> 'No, sir. Why is that? Because those things are not our self and do not pertain to our self.'
>
> 'Even so, monks, whatever is not yours: Let go of it. Your letting go of it will be for your long-term happiness and benefit. And what is not yours? Form (body) is not yours ... Feeling is not yours ... Perception ... Mental Processes ... Consciousness is not yours. Let go of it. Your letting go of it will be for your long-term happiness & benefit.'
>
> (MN 22)[5]

It is clear from this selection of suttas that the assumption or sense of a self, at least in relation to the *khandhās*, is something that Buddhism considers, like *taṇhā*, to lead to *dukkhā*. And, as with *taṇhā*, the sense of self is to be discarded if *dukkhā* is to be eliminated. The relation between the sense of self and *taṇhā* is thus very close – a point I return to later. We can also gather that the *self* Buddhism concerns itself with here is not lofty and esoteric, but is something that an average 'run-of-the-mill' person, harbouring *taṇhā*, will assume himself to be. The phrase 'run-of-the-mill person [who is] not well-versed or disciplined in their dhammā' applies to anyone who is not an *Arahant*, to anyone who still harbours *taṇhā* in their mindset. Such a person, harbouring *taṇhā* and a sense of self towards the *khandhās*, will be reborn into *saṇsārā* and *dukkhā*. On the Buddhist picture almost all of us will be such a person.

A further point to come across is that the self that a 'run-of-the-mill' person assumes he is, is some kind of *owner*. It is our sense of being a self *qua* owner that leads to *dukkhā*; the impression of owning *khandhās* is one that Buddhist practice will seek to eliminate. To be an owner, as MN. XXII indicates, is to stand in a relation of 'belongingness' to something else. From SN XXXV. 205 we can gather that a *sense of* belongingness involves thoughts of 'me', 'mine' or 'I am', while MN 22 later alludes to ownership through an injunction not to regard *khandhās* as 'yours' (hence not with thoughts of ownership) and SN XXII.99 alludes to the owner-self as 'possessing' the *khandhās*. What, then, does it mean to be an owner, such that one *possesses* the *khandhās*? Whatever it means, it will have to be distinguished from two other prevalent kinds of ownership that I shall for convenience term 'perspectival ownership' and 'possessive ownership'. I outline these contrast terms before saying more about the notion of ownership in the suttas.

For a subject to own something in a *perspectival* sense is for that thing, an object, to appear to the subject seemingly in a way that it can appear to no other subject.[6] All 'private' phenomena such as thoughts, intentions, perceptions and sensations – at least as they appear to a subject – will be

perspectivally owned by the subject; i.e., *I*, as a subject, view them from *my* perspective. In this sense, we commonly speak of such things as 'my toothache, my thought about Canada, my perception of the sea, my intention to move'. When it comes to objects (such as trees) that are agreed to be *external* to the 'inner life' of a person, what will be perspectivally owned is not the tree but the specific manner through which the tree or other object appears to an observing subject, by means of the relevant sensory input (visual, auditory and so forth). In relation to objects that appear to a subject in this special way – and which can hence be construed as 'mine' in this sense – the subject may be termed a 'perspectival owner'.

For a subject to own something in a possessive sense is for that object to be regarded as theirs by right of social convention, actual or ideal. Ownership of clothes, houses and even people can fall into this category. In relation to possessively owned objects, the subject can be termed a 'possessive owner'.

It seems plausible to suppose that a subject's ownership of its body involves, at the very least, both perspectival and possessive ownership. For on one hand, a person's 'own' bodily existence and movement (hence action) can be known directly to the associated subject 'from the inside' by way of perspectivally owned perceptions and sensations. In this context, the subject may speak more broadly of 'my body' or 'my actions' in a perspectival sense. On the other hand, one's body can be regarded as a publicly observable object that will one day die, or whose parts can be lost or donated while alive. In a context where one may donate blood as one donates clothes, the subject can be described as having or renouncing possessive ownership of it (or of part of it).

The notion of ownership alluded to in the Buddhist suttas seems neither perspectival nor possessive, but is nonetheless implicated in the prevalent attitude one has towards one's body and mind (a combination of various *khandhās*). Importantly, the *Arahant* is depicted as lacking any sense of this kind of ownership towards such *khandhās*, which immediately distinguishes it from the other two sorts of ownership. Since *Arahants* still interact with the (conditioned) world, they could not for instance lose the impression of perspectivally owning objects; objects, such as physical *dukkhā*, will still appear to them in a way that can be accessed by no other subject. Nor will they suddenly cease to recognise such social conventions as those that dictate possessive ownership of (say) their robe and bowl. While neither perspectival nor possessive, this third kind of ownership, a sense of which *Arahants* are depicted to lack, is clearly portrayed in the suttas as a sense of ownership had by *most people* towards various *khandhās* – in particular, as it happens, those 'very' *khandhās* that are perspectivally or possessively owned by them. Perhaps, then, it a kind of ownership-sense that a person could not decipher with introspective ease unless she had experienced its loss. When we put the point this way, we can immediately point to (although not yet

clearly define) a notion of ownership that is alluded to in contexts of Western psychology and which seems to answer to what is talked about in the suttas. It is a sense of ownership or 'my-ness', which, while ubiquitous, may on occasion be lost, suspended or compromised. Unlike with the *Arahants*, the loss of ownership-sense in such cases does not seem to affect their mindset in a global manner, and, rather than being beneficial, it is pathological. It involves the person regarding a subset of *khandhās*, formerly felt as theirs, to 'not belong' to them. I shall henceforth refer to this kind of ownership – which still requires further elucidation – as *personal ownership* whose subject identifies as a *personal owner*. A *sense* of personal ownership with respect to selective body parts or mental objects, then, is reportedly found lacking in cases of depersonalisation and anosognosia. Anosognosia is a mental affliction where the patient has a physical deficit (such as paralysis) but does not, on some cognitive level, recognise that he has it. As the philosopher Drakon Derek Nikolinakos reports, this can elicit a denial of (personal) ownership towards the affected body part:

> When such patients deny a deficit, they also tend to show indifference toward the affected body part or deny ownership of it and justify such denial with confabulations, e.g., 'the limb was left behind by another patient', 'the limb belongs to the examiner'. They may also ignore the request to move the affected limb, or they may respond to the request by saying 'here you are' without being able to acknowledge that no movement has taken place …. Patients who combine anosognosia and limb paralysis present a graded reflexive consciousness when they acknowledge the condition of paralysis in parts of the affected limb but deny it with respect to other parts of the same limb, e.g., the patient who denied ownership of the left hand but did not deny paralysis of the left arm and elbow … Such patients may persist in their denial of ownership of the specific body part in spite of the visual evidence, which may even be acknowledged explicitly. They may even acknowledge the oddity of their statements that it is paradoxical to hold that the forearm and hand, which they do not consider as belonging to them, are attached to their elbow.
>
> (2004, 323–324)

Now someone with a paralysed arm but *without* anosognosia will usually associate the visual cues of it as attached to their body – a body that seems perspectivally owned because of such cues – with a sense of personal ownership or 'my-ness' towards the body and its paralysed arm. In the above example, however, it does not seem as if the *perspectival* cues of seeing one's paralysed arm as attached to the body are being associated with the kind of *personal* my-ness that is usually felt towards one's arm, paralysed or not. That the patient is aware of the visual cues (and so has some sense of

perspectivally owning the arm) is evident through the fact that he can be puzzled by a mismatch between the visual evidence and the lack of felt ownership.

When it comes to the perceptions, sensations and thoughts themselves (as opposed to external body parts inferred through such sensations, etc.), a decoupling of a felt personal ownership from perspectival ownership towards them can be yet more evident. Cases of *depersonalisation* provide good evidence that while some sensations, perceptions or thoughts present a unique perspectival aspect to the subject, the subject feels as if the very same mental objects do not belong to it.[7] It is *this* kind of 'personal' ownership or 'belongingness', then – whose *felt* loss is evidenced through cases of anosognosia and depersonalisation – that seems relevant in the above suttas. It is a kind of ownership that seems taken for granted by most people, but is introspectively hard to decipher, and whose feeling can be lost upon occasion. (We say more about the difference between global versus partial loss of the sense of personal ownership in later chapters.)

Can more be said about this notion of personal ownership? It is notable that Western philosophy does not seem to have formalised a distinction between perspectival and personal ownership, the two sometimes being conflated despite their differences.[8] Perhaps this is unsurprising since the two usually occur together: the suttas themselves implying that where there is perspectival ownership towards aspects of the mind and body (viz., towards the *khandhās*), there is almost always a sense of personal ownership towards those aspects. While the concept of personal ownership does not seem fully articulated in Western philosophy, the suttas may provide some clue as to what is involved. The key to their account seems to be that personal ownership (unlike perspectival or possessive ownership) is to be analysed in terms of *identification* (although they do not use that exact term). In what follows, I will say more about what is meant by identification. Embedded in this analysis will be a definition of personal ownership in terms of identification. I will then provide evidence from the suttas to suggest that Buddhism defines the *self* in terms of personal ownership and hence, in terms of identification.

1.2. Identification and the self

Identification involves a reflexive assumption, on the part of the subject, that various (psycho-physical) attributes are in some way assimilated into itself. For purposes of our analysis it is useful to draw upon a further distinction between, on the one hand, the terms 'identify(ing) with' and 'identify(ing) as', which, in the verb form, depict what the subject implicitly *does* (its act of assuming itself to be assimilated with the attributes), and, on the other hand, the term 'identified with' which, in the adjective form, describes the state of affairs assumed to be true by the subject (itself as assimilated with various psycho-physical attributes). Hence we can say that by

identifying with or *identifying as* various attributes, the subject (reflexively) assumes itself to be *identified with* them in some integrative manner.[9] I will now say more about the distinction between the verbs 'identifies *with*' and 'identifies *as*', before expanding upon the adjectival 'identified with'.

I will, throughout this book, adopt the following convention of distinguishing between the verb phrases 'identifies with' and 'identifies as'. It should be seen as a useful stipulation rather than as a claim for any deep-seated existence of such a distinction. When a subject identifies itself *with* an item X, I stipulate that this implies (in our given context) an independent reality to the X that the subject identifies itself with (such Xs as the *khandhās*). When a subject identifies itself *as* X, on the other hand, then this does *not* imply the reality of X; X, in our given context, may or may not exist (X might be the self). 'Identifies *with* X' hence implies the reality of X, while 'identifies *as* X' does not imply the reality of X.

Now let us further consider the phrase 'S is identified with X' as it occurs in the above context. The term 'identified with' is to be construed either as X being assimilated with the whole subject, or as X being assimilated with a part of the subject, or as the subject being assimilated with X. It is important to note that the fact that a subject identifies itself with (or as) any X does not imply that X *really is* assimilated (viz., identified) with the subject in any way, anymore than a philosopher's 'arguing for Y' implies that the argument for Y succeeds.

With these distinctions in mind, we can now analyse the following very important excerpt from SN XXII.99 which alludes to the central relation between selfhood and identification. In this sutta, the 'run-of-the-mill person … (a) assumes [*khandhās* (form, feeling, perception, mental formations, consciousness)] to be the self, or (b) the self as possessing [*khandhās*], or (c) [*khandhās*] as in the self, or (d) the self as in [*khandhās*]'. I analyse (a)–(d) in turn.

(a) Suppose that a 'person' – which I think is most usefully analysed in this context as a *subject* – 'assumes' various *khandhās* – aspects of the associated body or mind – to 'be the self'. I take this to mean that the subject identifies itself with various aspects of the body or mind (or perhaps with all of them together), such that the aspect(s) somehow seems, from the subject's perspective, to be at one with the subject as a whole. Most basically, this will involve the subject – the *witnessing* as it presents from a psycho-physical perspective – identifying with those very *khandhās* (objects of awareness) that contribute, however subtly, to the impression of a hemmed-in perspective from which the world is witnessed. In Chapter 1, I mentioned that a person is likely to assume that witnessing and a particular psycho-physical perspective (together comprising the subject) are a psychologically basic unit: that witnessing must always present through such a perspective. On the Buddhist analysis, such an impression is actually the upshot of a primal identification on the part of witnessing-from-a-perspective. Through

identification, the witnessing assumes its outlook to be *intrinsically* confined to such a perspective, rather than, as an *Arahant* would view it, *extrinsically* confined (with the possibility of being free from it). 'The self', in this context, would thus refer to those perspective-lending *khandhās* as assimilated with the witnessing that shows through them.

What other groups of *khandhā* are likely targets of identification as a self? The body (viz., a *khandhā* of the type 'form') is a typical candidate. Suppose one thinks '*I* feel healthy today'. By identifying itself with the healthy body, the subject takes the body to be self, a single entity in which subject and body are fused. Common sense supports this analysis. When saying '*I* feel healthy' it does not seems right to analyse this as either '*the body* feels healthy' or '*the witnessing-from-a-perspective* feels healthy'. The 'I' that claims to feel healthy seems to be a hybrid, namely, the *body-as-subject*.

A subject might alternatively identify itself with *khandhās* of the type 'mental formation', perhaps through some assumed integration with an intellectual capacity, as borne out through such thoughts as '*I* am smart' or '*I* am stupid'. Again, this fits with common sense; for it does not seem right to analyse the statement as either 'the witnessing perspective is smart' or 'the intellectual capacity (or an idea of it) is smart'. The thing that purports to feel smart (or stupid) is rather something like the '*intellectual-capacity-as-subject*'. Through these reflexive assumptions, therefore, the subject will implicitly and reflexively feel *as if* various aspects of the body or mind are assimilated to itself in such a manner that those aspects *qua* subject are implicitly assumed to be a singular 'me' or 'self'.

While analysis of other suttas in this chapter will depict this identified-as self to have further features, we can already surmise from SN XXII.99 that the identified-as self will be implicitly presented, through a subject's act of identification, as something *gestalt* (in its seeming unification with the various *khandhās*), not *merely* as the haphazard addition of a subject to various *khandhās* (more on this later). Furthermore, it will seem to the subject as if the status of selfhood is not *created*, but *reflected* or *revealed* through an act of identification; although Buddhism contends, importantly, that the supposed self is *in fact* created through repeated acts of identification on the part of a subject (hence the alleged possibility of a subject being able to practice at *undoing* identification and the resultant sense of self).

When a subject 'assumes [various *khandhās*] to be the self' we can say, therefore, that a subject through identifying itself *with* various aspects of the body–mind, assumes itself to be more than just (the contingently confined) witnessing-from-a-perspective; it identifies itself, in this capacity, *as* a singular, integrated self or me. The 'self', as something that the subject identifies itself as being, minimally designates a subject to which aspects of the body–mind are somehow assimilated. 'Sense of self' (in this context) designates the *feeling/assumption*: 'I am this aspect of body–mind, this aspect of body–mind is what I am'.[10]

(b) When various *khandhās* are assumed to be the self, such an assumption will often be evidenced through a particular *concern* taken in the relation between 'subject *qua* bodily/mental self' and various *khandhās* that are owned in what I have stipulated to be a possessive or perspectival sense. This now relates to the part of the sutta that speaks of 'the self as possessing *khandhās*' – and from here a definition of personal ownership can be surmised. The sense of a <self *qua* body/mind-as-subject> (held on the part of the subject) will *give rise* to such assumptions as *personally owning* other *khandhās* in *relation to* this assumed identity. Suppose one buys a pair of trousers that fits very well – a fact that one delights in. One's prior identification as a bodily self now feeds a sense of my-ness towards the trousers that is suggestive of more than just *possessive* ownership. It is a sense of *personal* my-ness which reflects the subject's identification as a somebody who is the proud owner of the flattering trousers. In this capacity, a subject identifies *with* the possessive owner of the trousers and hence *as* their personal owner. More precisely, by identifying *with* those aspects of body–mind (*khandhās*) that feed into the idea of possessively owning the trousers – including the feelings of pride that arise from how the trousers appear on the pre-identified-as self – the subject identifies *as* the personal owner of the trousers. 'The self' in this context refers to *me as personal owner* of the trousers: the possessive owner as assimilated with the subject. Not only are the new trousers possessively owned; there is also, through feelings of pride in its possessive ownership, a sense of *personal* ownership or my-ness towards the trousers – and it is *this* kind of 'possession' to which the sutta seems to allude.

Take another example: suppose that one really dislikes the feelings of nervousness that arise before having to give a speech. Having already identified as a self by taking various aspects of the mind to be 'me', an intimate and personal sense of 'my-ness' is extended towards those nervous feelings that is suggestive of more than just perspectival ownership. By thinking 'I wish *I* wasn't feeling nervous' a subject implicitly identifies *with* those aspects of the mind that stand the subject in a relation of perspectival ownership towards the nerves and hence, in this capacity, *as* their personal owner.

Emerging from this analysis, then, is a definition of personal ownership in terms of identification. Whenever a subject has feelings of personal 'my-ness' towards any object, bodily or mental, then the subject implicitly *identifies with* whatever group of mental and bodily *khandhās* serve to stand it in a relation of possessive or perspectival ownership towards the object in question. (This analysis will be expanded upon in Chapter 4.)

(c) Suppose that the possessive trouser-owner now decides that her messy hair (in particular) is quite out of keeping with her smart image. The kind of attitude towards the messy hair in particular, may be borne of assuming '*khandhās* … in the self'. The subject does not merely view the hair as any old hair, but views the hair as a part of the *self*, that is, as an integral *part* of the body–mind-as-subject that 'I' want to change.

(d) After the new hairdo, we can suppose that one now reflects: 'Finally I fit into the corporation; I am properly a part of it'. This is assuming the 'self as in the *khandhās'*, where the subject is identifying itself with some further set of *khandhās*, those that make up the corporation.

While this account of identification requires further elaboration – a task postponed until Chapter 4 – we will have gained some initial idea of what is meant by the notion of 'identification' in relation to SN XXII.99. The kind of self that the subject *identifies itself as being* will, on the Buddhist account, minimally pick out a subject that is, in various capacities, identified (viz., assimilated) with various *khandhās*.

1.3. Personal ownership, the self and identification

In the following very important sutta (abridged), the self, which we can now surmise must involve identification, is further defined in terms of *personal ownership* such that whenever there is a self or 'me' (implying identification), then there must be what is personally mine, viz., what is *personally owned*; and conversely, whenever there is what is *personally owned* or mine, there must be a me, viz., a *self* or personal owner (implying identification). This sutta analytically links the notion of personal ownership to a self (and hence to identification), yielding a basic condition of what it is for an entity, on the Buddhist position, to be a self:

> 'Bhikkhus, there being a self, would there be for me what belongs to a self?'
> 'Yes, venerable sir.'
> 'Or there being what belongs to a self, would there be for me a self?'
> 'Yes, venerable sir.'
>
> (MN 22)

The primary candidates for 'what [personally] belongs to my self' are *khandhās* that are already owned in a perspectival or possessive sense. It is natural to interpret the second half of MN 22, first of all, to be saying that if something were to belong to a self – to be personally *owned* by a self – then the self would automatically assume the role of *personal owner* in relation to the owned item. The personal owner-self that the subject identifies itself as being would, given the analysis of SN XXII.99, amount to the subject as assimilated with all those *khandhās* that, at a time, underpin the roles of possessive or perspectival owner in relation to certain items. The first half of MN 22 seals the co-definition of 'self' with 'personal owner' (and hence 'personal ownership') by stating that if there is a self, then that self must stand in a genuine relation of personal ownership or my-ness to some other item, X. We have established (according to the second half of the sutta) that standing in a genuine (rather than merely apparent) relation of personal ownership or my-ness to some X will guarantee that the subject not only identifies *as*, but *is* (descriptively) identified *with* the self, a personal owner of X.

We can hence infer from our analyses of the sutta that (a) whenever there is personal ownership (or my-ness) towards some X, then there is a self (a me who is personal owner) *and* (b) whenever there is a self (a me), then there is a personal owner (via the personal ownership relation of my-ness to some further X). The role 'personal owner' thus emerges as basic to the notion of self in Buddhism; the self must, at the very least, be a personal owner. Of course, whether a subject *really is* identical to any such owner-self is a question at issue. I will henceforth refer to a subject's *genuine* identity with the self *qua* any specified role (such as *personal owner*) *as* 'self-identification' or 'self-identity'.

A personal owner, hence, involves a subject whose *identification with* (and personal ownership of) various *khandhās* reflects its overall self-identity with the role of personal owner. To have a *sense* of self, then, is (minimally) for a subject to reflexively feel as if it is identical to an owner-self. Should the subject turn out to *be* an owner-self, the sense of self must be grounded in an *owner-self* that is sensed. Should the subject turn out *not* to be an owner-self, then the sense of self will not be grounded in an *owner-self* that is sensed. When one has a *sense* of personal ownership towards X, then one has a *sense* of my-ness towards it, entailing a reciprocal *sense* of self-identification with the personal owner of X. We express this neutrally (in a way that does not presuppose the self's actual existence) by saying that when a subject identifies itself with various *khandhās* in a manner that reflects its assumed ownership of an object X, the subject *identifies as* the personal owner of X, namely *as* a self.

Applying this analysis to the anosognosic, we can now say that when he feels that his arm does not belong to him, this means that he fails to identify as the personal owner of the arm. In Buddhist terms he does not, with respect to the arm, view 'the self as possessing [*khandhās*]'. In failing to 'assimilate' to himself (*qua* subject) the idea of being the arm's perspectival owner (and so failing, through this lack of identification, to generate a sense of personal ownership towards his arm), he lacks a sense of self towards the arm; he does not feel as if he is the arm's personal owner. Or when he feels that the arm is not a part of who he is, this means that he (*qua* subject) fails to *identify with* that body part *as* a *part of himself* (*qua* self). In Buddhist parlance he does not, with respect to the arm, view 'the *khandhās* as in the self'; he does not assimilate the arm into the set of *khandhās* that seems to reflect his identity as an integrated self or 'me'. In Chapter 4, I further develop the notion of assumed self-identity, in particular, as it relates to the roles of perspectival and possessive ownership.

1.4. The sense of self (through reciprocal senses of personal ownership and self-identification) co-arises with *taṇhā*

It is clear that Buddhism regards a sense of a self to be something that leads to *dukkhā*. The ideal mindset, as suggested in SN XXXV. 205, is one where a

person regards the *khandhās* in such a way that 'thoughts of "me" or "mine" or "I am" do not occur to him'; they no longer play a role in his attitudes, motivations or choices. Given the analysis so far, we can surmise that a lack of 'mine' thoughts amounts, specifically, to the lack of a sense of personal ownership towards the *khandhās*, and the reciprocal lack of 'me' or 'I am' thoughts pertains to a subject's no longer identifying with any of the *khandhās* as *me* or as a *part of me*. Having renounced a sense of self-identification and personal ownership towards the *khandhās*, where one no longer regards them in this manner to be *me* or *mine*, one is also 'set loose' from *dukkhā* (mental *dukkhā* while physically alive, mental and physical *dukkhā* when the body dies). Now we already know from the Third Noble Truth that Buddhism regards a mindset with *taṇhā* to lead to *dukkhā*, and the cessation of *taṇhā* to lead to the cessation of *dukkhā*. Given that a parallel story is true with the sense of self in relation to *dukkhā* – in virtue of the subject's reciprocal senses of personal ownership and self-identity – we can surmise that the relation between *taṇhā* and the sense of self will be very close. The two will, at the very least, arise together.

For purposes of this project, I will treat the proposed relation between *taṇhā* and the sense of self as an empirical relation, such that (if the relation were to hold) whenever there is a sense of self, there is *taṇhā*, and whenever there is *taṇhā*, there is a sense of self. From our analysis, those aspects of self-sense most relevant to *taṇhā* will involve the subject's implicit assumption of being *identified with*, or a *personal owner* of, various *khandhās*. Analysing the relation this way opens up the possibility of its empirical investigation – a possibility to be explored later in the project. I anticipate that if the Buddhist analysis proves accurate – with the senses of self-identification and personal ownership co-arising with *taṇhā* and hence potential mental *dukkhā* – then it will contribute substantially to the discussion of ownership and identification in the West. For not only will it clarify the concepts of ownership and identification; it will offer an effective empirical means of *testing* whether a person is generally identifying as a personal owner – and hence as a *self* – and if so, the items towards which they are harbouring a sense of personal ownership or self-identity.

Suppose we limit the pool of potential items X (towards which one might have a sense of personal ownership or self-identification) to those that are perspectivally and/or possessively owned. The Buddhist claim would therefore be that there is *also* a sense of personal ownership or self-identification with respect to item X if and only if one has *taṇhā* with respect to X (where the bi-conditional is regarded as indicative of empirical co-dependence). Determining if one has *taṇhā* with respect to X might be achieved in a number of ways, most crudely, by imagining: 'If X were to be lost or gained, would my happiness or (mental) suffering alter?' If 'yes', then one would likely harbour *taṇhā* and hence a sense of personal my-ness or self-identification with regards to X; if 'no', then one may well not harbour *taṇhā* with respect

to that X. In the case of the anosognosic with the foreign-feeling arm, the Buddhist analysis would predict that he harbours little or no *taṇhā* towards it, and is therefore emotionally indifferent to its fate, feeling no negative emotions if it were to be harmed. It is interesting, therefore, to note that Nikolinakos (above) mentions *indifference*, alongside the lack of ownership-sense, to be a common feature of anosognosia. Should the anosognosic suddenly become emotionally concerned again about the arm's welfare, our Buddhist analysis would predict a return of *taṇhā* towards it and hence a return of his sense of personal ownership towards the arm – with the reciprocal sense of self-identification as its personal owner – and, quite probably, an elimination of the anosognosia with respect to that limb.

Where an item of *taṇhā* is neither possessively nor perspectivally owned, such as a political situation, the right thing to say will probably be that one identifies strongly as the *personal owner* of their thoughts, opinions and feelings that relate to that item. There may well be further notions of (non-personal) ownership to which a sense of personal ownership might be annexed.

1.5. Further features Buddhism ascribes to the reflexively assumed self

I have spent some time analysing the Buddhist notion of (commonly assumed) self as centrally a *personal owner* – of which we can take as given, its attitude of reciprocal personal ownership towards the *khandhās*. I now consider further features that Buddhism would ascribe to the self. We can immediately note that built into the role of 'personal owner' is the central but implicit aspect of *witnessing-from-a-perspective* (which identifies itself *as* a personal owner). 'Subject' can thus be included in the definition of self extracted from the Buddhist suttas.

The feature of *synchronic unity* (unity at a time) has been tacitly alluded to in connection with the subject's assumed *assimilation* with various *khandhās*. Underpinning the subject/object division in Chapter 1 was the observation that witnessing-from-a-perspective seems to escape, in principle, the purview of attention (reflexive or otherwise). It stands to reason, therefore, that if a subject identifies itself with various *khandhās* at a time, such that the *khandhās seem* assimilated to the subject (as 'self'), then those *khandhās* will *also* appear to elude the subject's attentive radar. From the subject's viewpoint, the *khandhās* will not overtly appear as separate *objects*. If the *khandhās* did overtly appear to the subject as items amenable to attention, then how could they possibly *seem* identified with the witnessing (unobject-like) *subject* in the first place? In view of this, it makes sense to suppose, on the Buddhist analysis, that a subject's identifying itself with various *khandhās* at a time will be felt as a unity: the *khandhās* seeming somehow integral to – at one with – the subject who identifies with them as (or as part of) a self. We can hence ascribe the notion of 'synchronic unity' to the self of Buddhist analysis.

The subject's assumed unified identity with various *khandhās* will, given this analysis, exert a curious 'blinding' effect upon the way that those *khandhās* (viz., objects of awareness) are apprehended – or more fittingly, upon the way they *fail* to be apprehended (at least attentively). We have noted that Buddhism regards the objects of the world – parsed into the psycho-physical *khandhās*[11] – to be conditioned and hence impermanent (*aniccā*, pronounced 'annie-cha'). Now if various *khandhās*, by being identified with by a subject, fail to be overtly noticed by the subject as *objects* of witnessing, then we can infer that they will also fail to be overtly noticed as *aniccā* – as obviously coming and going from the subject's (most likely peripheral) field of witnessing. It will be as if the subject, while assuming various *khandhās* to be united with itself *qua* self, becomes somehow change-blind to their coming and going, even though their presence in the field of witnessing will be inattentively felt. The skewed perception of their impermanence will also imply, on the Buddhist analysis, that their status as causing *dukkhā* when emotionally invested in (or as not being conducive to *dukkhā*-free happiness) will fail to be apprehended with clarity. Buddhism maintains that if one were to clearly apprehend the *khandhās* as *aniccā* and therefore as *dukkhā*, then there would be no incentive to emotionally invest in them (and hence to regard them with *taṇhā*). That we do emotionally invest in the *khandhās*, according to Buddhism, reflects our lack of clear apprehension of their nature as *aniccā* and *dukkhā*.

This inference about the shortfall in apprehending various *khandhās* as *aniccā* and *dukkhā* – due to a subject's sense of self-identification – is borne out in the oft-cited *Anattā-lakkhaṇa* Sutta (SN XXII.59), where the Buddha urges his followers to regard the *khandhās* as *aniccā*, *dukkhā* and *anattā*. To fully regard the *khandhās* as *anattā*, we can now surmise, is to regard them as not pertaining to a self, in particular, as stripped of the reciprocal senses of self-identification and personal ownership. While it is *these* aspects to the self-sense that Buddhism regards as being linked to *taṇhā* and *dukkhā* – and hence to be wisely eliminated – there is *more* to the Buddhist notion of self than just a personal and unified witness-owner. The *Anattā-lakkhaṇa* sutta, besides bearing out the above inference, contains further central clues as to how Buddhism construes the self that we have a sense of being:

> I have heard that on one occasion the Master was staying at Varanasi, in the Game Refuge at Isipatana. There he addressed the group of five monks:
>
> 'Physical form, monks, is not the self. If physical form were the self, this body would not lend itself to dis-ease. One could get physical form to be like this and not be like that. But precisely because physical form is not the self, it lends itself to dis-ease. And one cannot get physical form to be like this and not be like that.
> 'Feeling is not the self ... Perception is not the self ... Mental processes are not the self ...

> 'Consciousness is not the self. If consciousness were the self, this consciousness would not lend itself to dis-ease. One could get consciousness to be like this and not be like that. But precisely because consciousness is not the self, it lends itself to dis-ease. And one cannot get consciousness to be like this and not be like that.
> 'How do you construe thus, monks – Is physical form constant or inconstant?' – 'Inconstant, Lord.' – 'And whatever is inconstant: Is it easeful or stressful?' – 'Stressful, Lord.' – 'And is it right to assume with regard to whatever is inconstant, stressful, subject to change, that 'This is mine. This is my self. This is what I am'?' – 'No, Lord.'
>
> '... Is feeling constant or inconstant? ... Is perception constant or inconstant? ... Are mental processes constant or inconstant?...
> 'Is consciousness constant or inconstant?' – 'Inconstant, Lord.' – 'And whatever is inconstant: Is it easeful or stressful?' – 'Stressful, Lord.' – 'And is it right to assume with regard to whatever is inconstant, stressful, subject to change, that 'This is mine. This is my self. This is what I am'...' – 'No, Lord.'
>
> 'Thus, monks, any physical form whatsoever that is past, future, or present; internal or external; blatant or subtle, common or sublime, far or near: every physical form – is to be seen as it actually is with right discernment as: 'This is not mine. This is not my self. This is not what I am.'
> 'Any feeling whatsoever ... Any perception whatsoever ... Any mental processes whatsoever ...
> 'Any consciousness whatsoever that is past, future, or present; internal or external; blatant or subtle, common or sublime, far or near: every consciousness – is to be seen as it actually is with right discernment as: 'This is not mine. This is not my self. This is not what I am.'
>
> (SN XXII.59)[12]

From the imperative tone of this sutta, we can surmise that the Buddha would not be urging people to perceive the *khandhās* as impermanent (*aniccā*), *dukkhā* and without self (*anattā*) unless there was an *existing psychological tendency* to regard them as otherwise. So what is most likely to account for such a tendency? From SN XXII.59 we can gather that the tendency to not view *khandhās* as *aniccā* and *dukkhā* has to do with viewing them in terms of a self; the sense of self, through identification, is somehow obscuring our perception of the *khandhās* as *aniccā* and *dukkhā*. Could there be something about their identification as a *self* that makes it seem as if the *khandhās* lack *aniccā* and *dukkhā*? Indeed, could there be something about the assumed self that *itself* seems to lack *aniccā* and *dukkhā*? The sutta provides a strong hint that the commonly assumed self does, in some important way, lack these characteristics.

'Dis-ease' (meaning a lack of perfect ease) and 'stress' are Thanissaro Bhikkhu's translations of *dukkhā*. When the Buddha says of each of the

khandhās that they are not a self because they lend themselves to *dukkhā*, the implication, spelt out above and noted by Harvey (1995, 46), is that *if* some item X were a self, *then* that X would not lend itself to *dukkhā*. Not-*dukkhā* can hence be surmised as being attributed, in some way, to the self that we assume we are (and we shall see that how 'in some way' gets spelt out is crucial to the analysis). A parallel story can be applied to *aniccā*. The sutta suggests that it is not right to regard what is inconstant, stressful (*dukkhā*) or subject to change as a self. The implication, also noted by Harvey (1995, 46), is that if some X were to be rightfully assumed as a self, then that X would not be inconstant or subject to change – not *aniccā*, in other words. The implication is also, crucially, that most of us *do* take ourselves to be an entity that somehow lacks in *aniccā* and *dukkhā*; an assumption that warps our perception of various *khandhās* into being viewed as 'mine', 'myself' and 'what I am', and hence, as somehow lacking in *aniccā* and *dukkhā*. Finally, as Harvey also notes, the opening passage of SN XXII.59 suggests that if any of the *khandhās* were a self, then they would be under one's control, and hence not lend themselves to *dukkhā* (on the assumption that we seek to avoid *dukkhā*). Harvey (1995, 49–50) suggests (with backing from further suttas) that the notion of control in this context be construed reflexively, implying that the controlled self is under its own control, making the self a controller or agent.

Now, when using this derived information such as 'not *dukkhā'* and 'not *aniccā'* to help define that self-entity alluded to in Buddhism, we need to be careful. The temptation is to confidently ascribe the positive counterpart of these attributes to the self-entity, as if *this* is the kind of entity that Buddhism is mainly talking about. Hence Harvey concludes his (otherwise insightful) analysis of the *Anattā-lakkhaṇa* and other suttas with the statement:

> It can thus be seen that the Self-ideal which early Buddhism worked with was of an unconditioned, permanent, totally happy 'I', which is self-aware, in total control of itself, a truly autonomous agent, with an inherent substantial essence, the true nature of an individual person.
>
> (1995, 51)

The self depicted here sounds like a close cousin of nibbānic consciousness! Harvey indeed notes the explicit similarity with *nibbāna*, stating the only differences to be that the self, unlike *nibbāna*, carries the identity of 'I am' or 'this I am' as well that of being a controller. Buddhism does not, of course, construe the nibbānic element as the 'agent-controller of action, as the Self is seen to be' (Harvey, 1995, 53). *Nibbāna* is also not to be construed as an individual, bounded entity: a corollary of it not being a locus of self-identity.

While I believe Harvey's comparison between the assumed self and *nibbāna* to be on the right track (many scholars do not acknowledge any connection), it clearly needs refinement, since any self we assume we

are, even if in some sense lacking *aniccā* and *dukkhā*, will *not* be manifestly an unconditioned, permanent, totally happy 'I'. We only need to reflect upon the amount of misery in the world, perhaps in our own minds, to see that an assumption of our self-nature as perfectly happy is not common amongst humanity. We only need to reflect upon the widespread fear of death and non-existence to see that an assumption of being unconditioned or permanent is not ubiquitous. This does not rule out the possibility that somehow, underneath all the *dukkhā*, the nature of the mind is that of perfect, unconditioned, nibbānic happiness. But if the nature of mind is perfect in this way, then it is not glaringly obvious. It is indeed unlikely that the Buddha would have set forth a Noble Eightfold Path to the cessation of *dukkhā* if he had thought that our assumed self-nature was already perfectly happy, as Harvey's analysis would have us believe.

Now this is not to deny that the Buddha, in other passages, *did* caution against propounding notions of a self that closely matched the description of Harvey's passage. But we need to understand the context in which such cautions were made. Around the time of the Buddha, it was apparently common for scholars and pandits to expound various theoretical views of a self, views the Buddha regarded as leading to *dukkhā* if dogmatically adhered to. The Buddha's injunction for pandits and scholars to renounce such views of the self should not distract us from the main force of his suttas on *anattā*: to lose that sense of self that almost all of us – theoreticians or not – are said to harbour. Our current aim is to get a plausible picture of how Buddhism would construe this commonly assumed self, a picture that does not make it sound obviously like *nibbāna*, but which may yet include features that, given Buddhist metaphysics, are most parsimoniously explained by nibbānic consciousness.

From SN XXII.59, we have clues that the assumed self is not *dukkhā*, not *aniccā*. I think that the most intelligible way to construe the self (on the Buddhist account) as being 'not *dukkhā*' is, perhaps paradoxically, through looking to *taṇhā* (the root of dukkha!). *Taṇhā* involves a misdirected urge to seek perfect happiness and avoid *dukkhā* – misdirected, since the urge is sought to be satisfied through seeking ideal configurations of the impermanent *khandhās* (which leads to *dukkhā*). Misdirectedness aside, on the (Buddhist) assumption that the urge to seek happiness and avoid *dukkhā* is itself universal among humankind, it makes the most sense to construe the 'not-*dukkhā*' element of self as simply *that aspect which seeks happiness and avoids dukkhā*. When combined with the assumption that this urge can be properly satisfied through an ideal configuration of *khandhās* (an assumption borne out through not properly apprehending their impermanence) it is only natural to suppose that the subject will fail to properly construe such an aim as conducive to *dukkhā* – hence prompting the Buddha's imperative in SN XXII.59 to correct this bias by regarding the *khandhās* as *dukkhā*.

Thanissaro Bhikkhu's translation of SN XXII.59 conveys *aniccā* or impermanence as having two aspects to it, or at least as being describable in two ways, namely through *inconstance* and *being subject to change*. I take 'inconstance' to mean a lack of numerical continuity, such that any conditioned object or its fleeting parts are prone to decay. I take 'subject to change' to mean that even if a conditioned object is conventionally designated as existing numerically over a period of time, its qualities, in virtue of its fleeting parts, will change or vary from one moment to the next. Now SN XXII.59 hints that the self we assume we are is somehow construed as lacking in *aniccā* and hence, as lacking in inconstancy and changeability. And I suggest that there is a way of construing the self, in this manner, which avoids an implausible casting as death-defiantly constant and unchangeable. The suggestion is as follows: from the viewpoint of the apparent self (which is actually the viewpoint of the witnessing subject that assumes it is a self) it is as if, from one conscious moment to the next, the self's numerical existence is not gappy or perduring (meaning with different temporal parts). *From one conscious moment to the next, there is a sense of the self's unbroken presence.* The qualitative invariability can be construed in a similar way: from the viewpoint of the apparent self*, it is as if there is something essential about the self – its very me-making quality – which does not, from one conscious moment to the next, change in quality. It stays exactly the same*. This can seem to hold true despite the changing flux of *khandhās* that are over time being identified with as 'me' or as 'a part of me'. For it can be supposed that such identification will seem, from the outset, to be effected by not a mere subject but by a qualitatively invariable independent self.

Now I am not of course claiming at this stage that this is indeed how we construe our selves; this argument is left to Chapter 4. The current claim is rather that such a construal makes good sense of how Buddhism would characterise the self that we assume we are. On one hand, the construal manages to avoid the obvious implausibility of the self seeming to defy death or non-existence, while on the other hand it bestows the assumed self with enough permanence to plausibly interfere with one's perception of identified-with *khandhās* as impermanent (inconstant and variable). For if we assume ourselves to be an entity whose existence is uninterrupted and does not change (from one conscious moment to the next), then it is natural to suppose that *while* the subject assumes various *khandhās* to be assimilated with this self-entity, namely, as 'me' or 'part of who I am', those *khandhās* will, from the subject's perspective, lose their appearance as impermanent. Fear of death may be explained as an oscillation between viewing the body as 'me' and viewing it as an object that will one day perish. Buddhism makes no claim that the things a subject identifies with as 'me' will not *in fact* vary from one moment to the next.

1.6. What nibbānic consciousness could bring to the sense of self

Our analysis of Buddhism, thus far, has unearthed the following notion of self: a subject that is a personal owner and agent, who seeks happiness and avoids *dukkhā* and who is unified and with an unbroken invariable presence from one conscious moment to the next. The sense (or assumption) of self, on the Buddhist analysis, is accordingly one's subjective impression of being such an entity that includes these features.

I have presented evidence from the suttas to suggest that the sense of being a personal owner is closely tied, on the Buddhist account, to a subject's tendency to *identify* with the *kandhas* as 'me' – a *taṇhā*-related tendency that Buddhist practice aims to eradicate. Our question is now: given what we know about Buddhist metaphysics, what is most likely to account for the other features ascribed to the self: synchronic unity, unbroken invariable presence and a happiness-seeking urge? In what follows, it will be argued that it is most economical to deploy *nibbānic consciousness* to help explain, on Buddhist analysis, our sense of these features in the self. Combined with the argument in Chapter 2, that nibbānic consciousness manifests through the witnessing element of the consciousness *khandhā*, it will follow that unity, unbrokenness (etc.) are carried into the self-sense via the witnessing element of this *khandhā*. Considered in its relational and object-oriented capacity, however, the consciousness *khandhā* will continue to be treated (along with the other *khandhās*) as *aniccā* and *dukkhā*. I will suggest that, construed as *aniccā* and *dukkhā*, the *khandhās* will not by themselves offer the simplest explanation for the impression of unity and so forth. They will instead be involved, through processes of identification and *taṇhā*, in 'watering down' the presentation of those features – in making nibbānic consciousness appear more like an ordinary self. The presence of *taṇhā* (etc.) will thus modify the effect of unified, unbroken (etc.) nibbānic consciousness to such an extent that the self comes across as something that we could reasonably identify ourselves as being, a hybrid entity perhaps, with features merging conditioned with unconditioned.

The urge to seek happiness and avoid *dukkhā* – which would plausibly contribute to one's sense of agency (we do not generally act unless there is some prospect, however minute, of personal gain in happiness or avoidance of suffering) – is perhaps most overtly influenced by *taṇhā*. Indeed, the urge to seek happiness (and avoid *dukkhā*) from the *khandhās* is partly what it *means* to have *taṇhā*. Yet on the Buddhist picture, it is still reasonable to ask: from where does our very intuition or belief in the ideal of *dukkhā*-free happiness arise? For if it is accepted that we perpetually seek *dukkhā*-free happiness from the *khandhās* (a claim implied by Buddhist psychology), then our very seeking of it would seem to suggest that we harbour an implicit hope and therefore belief that such happiness is possible to attain. What could account for this belief, given that emotional investment in the impermanent *khandhās*, on the Buddhist picture, could *never deliver* this ideal?

While it may be possible to explain this belief (in the ideal of *dukkhā*-free happiness) within a framework of the impermanent *khandhās*, it will have to be an explanation on which the belief's content has no grounding in reality; like a chased rainbow, the ideal will be thoroughly illusory. But another explanation is available in Buddhism. The ever-present 'underlying' nibbānic consciousness – which equates such consciousness to the pre-nibbānic witness-consciousness – can serve to explain, at least partially, our belief in this happiness-ideal. *Nibbāna* is after all depicted as 'not *dukkhā*', as the ultimate happiness in which it is possible to partake. But due to *taṇhā*, this real possibility of a happiness-ideal gets distorted into the belief that such an ideal can be found *outside* the intrinsic mind, in the perfect configuration of *khandhās* (in getting the world the way we like it!). As an urge perpetually frustrated (due to shiftiness of the *khandhās*) such *taṇhā* will be conducive to *dukkhā*; the ideal will never be realised. Yet allowing there to *be* the genuine potential for *dukkhā*-free happiness – in pre-nibbānic consciousness – involves having to tell less of a story about illusions; the explanation for our belief in such happiness is hence simpler than that which would involve just the conditioned *khandhās*.

Consider now the features of unbroken presence and invariability that are ascribed to the self. The subject's identifying with *khandhās* as a unified *self* (thus masking the impression of such *khandhās* as impermanent) will reinforce an impression of the *self's* unbroken presence and invariability. Identification will thus exert some influence on the appearance of *khandhās* as lacking in their native impermanence. Yet one can still reasonably ask, on the Buddhist picture: what best accounts for the very impression of unbroken presence and invariability, as it comes across through identification? Construed as impermanent entities, the *khandhās* seem intrinsically unsuited to deliver such an impression. It may be highly significant that Hume, with his empirically based ontology of impermanence, admitted defeat when it came to explaining the sense of what he called 'uninterruptedness' and 'invariability' of the self.[13] Any account that looks for an answer in the ontology of impermanence will be forced to conclude (in a similar way to before) that the features of unbroken presence and invariability are thoroughly illusory – and with this will come the burden of having to explain how those illusions come about.

As with the case before, Buddhism has an alternative answer to this problem, one that avoids the Humean burden of having to explain such a thoroughgoing illusion. Against the background of Buddhist metaphysics, the most economical and obvious answer will involve an appeal to nibbānic consciousness. From Chapter 2 it was determined that nibbānic consciousness, being timeless, cannot be *aniccā* (gappy and changeable) – just as, partaking in the nature of *dukkhā*-free happiness, it cannot be *dukkhā*. It therefore makes the most parsimonious sense to attribute the origin of the 'seamless invariability' impression, which *seems* inherent to a self, to the

witnessing aspect of a subject, namely, to pre-nibbānic consciousness. In this picture, the nibbānic witnessing, through the subject, and as part of consciousness *khandhā* (even though consciousness-of-an-object is impermanent), brings to the sense of self not only an impression of witnessing, or an implicit belief in the possibility of *dukkhā*-free happiness, but also an impression of unbroken presence and invariability.

As with the impression of a happiness-ideal, the presentation of unbrokenness and invariability, in pre-nibbānic consciousness, will be filtered through the distorting lens of identification. Not only will it seem as if a *personalised self* is the subject of such features, but it will also seem as if this self is caught up in time, with unbroken identity for the span of a life (as well as more immediately from one moment to the next). This will occur, for instance, when the subject identifies itself with a person of ten years ago, through taking a *khandhā* of memory (*qua* its content) to be 'self'. Including longer-term endurance and invariability is perfectly compatible with the *Anattā-lakkhaṇa Sutta* (SN XXII.59) making more realistic the Buddhist notion of self (most notions of the self include identity over time). And, as with immediate unbrokenness and invariability, it makes most economical sense to ground the core impression of 'the self's' longer-term *unchangingness* in the timelessly ever-present *nibbāna*. It is just that the more *temporal* dimension of *longer-term* endurance (etc.) will involve a greater degree of distortion from the identified-with and time-bound *khandhās* than will the more immediate sense of unbroken presence.

When a subject identifies with various *khandhās* as a part of itself, the *khandhās* are somehow felt, at any one time, to be unified with the identifying subject (and hence not viewed as overtly objectlike). The resultant impression of self will thus reflexively convey, at a given time, a *synchronically unified* subject-entity. Now we can ask, as we did in the case of unbroken presence: just what, on the Buddhist picture, will best account for this impression of a *unified* self? The identified-with *khandhās* will likely reinforce an impression of distinctive *me-ness* to this unified entity. The impermanent, differentiated *khandhās* do not however seem to offer an obvious explanation for the impression of unity itself (and we can note that Hume also admitted a failure to explain synchronic unity which he termed 'simplicity'). Nibbānic consciousness, on the other hand, has already been pegged (in Chapter 2) as undifferentiated and hence unified, providing Buddhism with an obvious answer to the origin of the unity-impression. The 'synchronic' unity will, on this picture, be inherent to the very witnessing that seems, like an in(di)visibility cloak, to enfold its unity around the identified-with *khandhās*.

We may finally wonder whether the 'unconditioned by relation' aspect of nibbānic consciousness will be likely to reveal itself at the pre-nibbānic level. Is there any way in which the self we assume we are, on the Buddhist picture, could present as being somehow apart from, and not dependent

upon, the *khandhās*? If there is, then such independence of 'self' from objects would not come across as overtly objective, for example, in such a manner that would render it eternal or non-physical. I suggest, nevertheless, that the witnessing subject *qua* self could still harbour a sense of ontological independence that comes through in the very impression of its being a kind of subject rather than object. The impression could plausibly be of a subject-self whose ontology has no contribution from any of the (attendable) objects *within its conscious purview*: objects that can be perspectivally owned by it. So the impression will be of an underlying self in charge of its mental retinue (such as thoughts and perceptions) rather than any of these attendable objects underlying *its* existence as a self. The impression would not be lost during identification. For when the subject identifies itself with *khandhās* as 'me' or as 'a part of me' (often through appropriating ideas of them), their statuses as separate, impermanent *objects* will, for reasons already mentioned, become effectively invisible to the subject. It would therefore implicitly appear to the subject (under its identification as a self) as if the self is, ontologically speaking, quite independent of any such objects *qua* attendable objects.

Now we are not saying that should the subject-self turn out to be real, it would in fact *ontologically depend* upon the *khandhās* (or ideas of them) that it identifies itself with – at least in their capacity as *objects*. To the contrary (and this point will be made clearer in Chapter 4), it will seem, to the subject, as if its essential selfhood stands ontologically prior to any such objects. It will seem to the subject as if a *pre-existing self* identifies itself with various *khandhās* (their object-like statuses rendered invisible), or as if a pre-existing identity of the subject with various *khandhās* is *revealed* on various occasions (as different aspects of 'me'), rather than as if the *khandhās* – in their capacity as impermanent objects – are helping to *construct* the self upon each new identification. On this picture – suitably attributed to Buddhism – identification will seem, from the viewpoint of the subject, to *evidence* rather than to *construct* an owner-self. Should the impermanent *khandhās* (such as thoughts or perceptions) *actually* help to construct the self, as Buddhism also contends, then that would undermine the self's reality, since the entity would lack an essential property of selfhood, viz., its ontological independence from any perspectivally ownable objects.

Hence, while the self will not reflexively present itself as being unconditioned by relation in an *objective* (e.g., non-physical) sense, we can nevertheless glean a way in which its *subject*ive existence could plausibly seem unconstructed by certain objects of witnessing. This feature, to be referred to henceforth as 'unconstructedness', will once again seem more difficult to accommodate within an ontology that admits only of co-dependent *khandhās*. For as conditioned things, each (tokening) of the five *khandhās*, on Buddhist cosmology, will manifestly depend upon other *khandhās*. The sense of the self's independence from certain objects of witnessing, insofar as it

seems *unconstructed* by them, will therefore be most parsimoniously cast as originating, on the Buddhist picture, in the pre-nibbānic consciousness whose intrinsic nature is to be unconditioned by relation. The conditioned nature of the identified-with *khandhās* will serve to modify the original 'nibbānic' contribution such that the 'self' comes across as being unconditioned in only the subjective (rather than objective) sense.

1.7. The definition and status of self in Buddhism

We are almost ready to offer a definition of that self which Buddhism (arguably) claims most people to have a sense of being. Before doing this, a further point needs mentioning, one that will later prove central to our analysis of the self. When the subject assumes itself to be a personal owner-self, through identification with *khandhās*, it crucially assumes itself to be a unified entity *with a boundary*. While boundedness (where 'the self' ends and 'the other' begins) will not be directly observable by the subject *qua* 'personal owner', it will be clearly *evidenced* through the fact that some *khandhās* are identified with as 'me', while other *khandhās* are viewed as 'not me'. Boundedness underscores a self/other distinction: the casting of self as an ontologically distinct and personalised entity. While the relationship between boundedness and self-identification will be spelt out further in Chapter 4, we can surmise that boundedness will be a feature that, as an upshot of a subject's identification as 'personal owner', is regarded in Buddhism to be on par with the role of *personal owner* (in similar relations to *taṇhā*). Without further ado, here, then, is the sought-after definition:

> A self is defined as a bounded, happiness-seeking/*dukkhā*-avoiding (witnessing) subject that is a personal owner and controlling agent, and which is unified and unconstructed, with unbroken and invariable presence from one moment to the next, as well as with longer-term endurance and invariability.

I surmise that *this* is the entity that Buddhism supposes that we – or most of us – assume we are. Is such an entity held to exist in Buddhism? The *assumption* or *sense* of being such an entity is certainly held to exist; it is after all the thing that, along with co-arising *taṇhā*, is deemed to perpetuate *dukkhā* – something to be erased, along with *taṇhā*, if *dukkhā* is to cease. But the entity itself – the self that we reflexively think of as inherent to our nature – has no place in Buddhist ontology; it cannot therefore be the *self* that is erased through the practice of the Noble Eightfold Path. The non-reality of self is perhaps most clearly expressed in the following oft-cited verses, where '*dhammās*' refer to all of reality, both conditioned and unconditioned:

> 'All conditioned things are *aniccā* (inconstant, changeable)'
> When one sees this with discernment

And grows disenchanted with *dukkhā*,
This is the path to purity.

'All conditioned things are *dukkhā* (stressful)'
When one sees with discernment
And grows disenchanted with *dukkhā*
This is the path to purity.

'All *dhammās* are *anattā* (not-self)'
When one sees with discernment
And grows disenchanted with *dukkhā*
This is the path to purity.

(Dhp XX.277–279)[14]

The first two verses allude only to conditioned things, namely the *khandhās*, and enough background in Buddhism has been presented for us to interpret their gist. Of current relevance is the third verse: 'All *dhammās* are *anattā*'. Armed with our analysis of what Buddhism means by 'self', we can inject some meaning into the inference that Buddhist ontology, while having room for that which is not *aniccā* or *dukkhā* (namely nibbānic consciousness), does not allow for the hybrid entity of self. To see all *dhammās* as *anattā* is to fully apprehend, reflexively and non-reflexively, that *everything* is not self (i.e., not personally 'me' or 'mine'): the *khandhās* and nibbānic consciousness alike. It is not as if losing the propensity to see the conditioned *khandhās* as (personally) 'me' or 'mine' implies a fallback position of seeing *nibbāna* as 'me' or 'mine'. In Buddhism, the impression of *me* or *mine* with regard to *anything at all* not only leads to *dukkhā*, but involves a fundamental misapprehension of reality, since there is nothing in the entire cosmological system, conditioned or unconditioned, that answers to 'me' or 'mine'.

Given our analysis, it is of utmost importance to understand that the Buddhist rejection of this self-entity from their ontology does not imply their equal denial of reality to each and every feature ascribed to the self. From the account presented, it should be fairly obvious that, *to the extent that they are not distorted by identification and taṇhā*, those features of witnessing, unity, unbroken presence, endurance, invariability, unconstructedness and the sought-after happiness-ideal, will *not* be considered to lack reality. It is most parsimonious, on the Buddhist picture, to suppose that the ground of these features is imported into the self-sense by the ever-present nibbānic consciousness (via witnessing). The features will only lack reality to the extent that they are unwittingly and reflexively ascribed to a *self* and are somewhat distorted (in their impression) through this process of ascription (viz., identification). The lack of the self's reality will pertain chiefly to the subject's purported status of being a bounded personal owner and agent whose independent existence *seems* to be reflected through the ubiquitous impressions of 'me' and 'mine' (including 'me' as the agent of actions). On the Buddhist

account, moreover, whenever the subject identifies as a bounded, personal owner-self, there must be *taṇhā* in the subject's mindset; simply put, *taṇhā* is present if and only if a sense of self is present. The upshot of this Buddhist analysis is that the sense of self, while not grounded in a real self, is contributed to by elements of which some have reality (being sourced in nibbānic consciousness) while some do not (being distorted by *taṇhā*).

Given this picture, the purpose of the Noble Eightfold Path of practice can be understood, in Buddhism, to selectively whittle away those tendencies that have the intrinsic nibbānic mind – with the nature of percipience – somehow 'duped' into assuming it is a bounded self-entity. What is whittled away will *not* be the witnessing with its contributing sense of unity and unbroken presence (etc.), but only those *taṇhā*-involving tendencies that have the subject assuming it is a bounded personal owner and agent *with* these features. Only when those tendencies to view the *khandhās* as 'me' and 'mine' have been completely erased will it be known with full discernment that 'all conditioned things are *aniccā*', 'all conditioned things are *dukkhā*' and that 'all *dhammās* are *anattā*'. This completes our analysis of how Buddhism is likely to define the ordinary sense of self and how its ontology, while having no room for that self we take ourselves to be, still acknowledges the contribution that *nibbāna* must make to the sense of self.

2. The misportrayal of Buddhism as endorsing a 'bundle theory' of persons

The analysis of Buddhism presented in Chapters 2 and 3 is rather unorthodox in its construing of the self as something unreal on the whole but with several features rooted in *nibbāna*. The orthodox position presents the self in Buddhism to be illusory in *all* its main features. In this section, I offer some reasons for why I think the orthodox position goes astray, along with illustrating examples of the orthodox position.

On a general level, one such reason may be that the self alluded to in Buddhism, whose existence is denied, has not been properly defined – either in the suttas or in secondary literature. Another reason may lie in that tendency for scholars to focus on the *anattā* suttas without considering their relation to the suttas on *nibbāna* in broader context of the Four Noble Truths. A further reason, which I have talked about elsewhere, could be that there seems to be an almost religious (and I think unwarranted) urge by many scholars to separate Buddhist from Upaniṣadic teachings.[15] Perhaps these reasons are not unconnected; the suttas on *nibbāna* make it more difficult to assert a principled distinction between Upaniṣadic and Buddhist teachings – so their significance is overlooked.

On a more specific level, if my interpretation is correct and witnessing is an element of the consciousness *khandhā* (as parsimony would seem to demand), then its truth, although consistent with construing the

'consciousness' *khandhā* as impermanent (due to its object-orientation), will not be obvious in the absence of analysis. For reasons mentioned before (of being pragmatically rather than ontologically oriented) the suttas only *imply* a notion of witnessing that is related to the impermanent consciousness *khandhā*; the concept of witnessing is not made explicit. The temptation will therefore be to view the ontology of persons,[16] on the Buddhist picture, as *thoroughly conditioned*: namely, as a bundle of thoroughly impermanent, *dukkhā*-ridden aggregates, *sans* any principle of unbroken percipience. It is against this background that Buddhism gets cast in the familiar orthodox reading of denying independent reality not only to a self, but also to its ascribed features that belie an ontology of impermanence: unity, uninterruptedness, invariability and that which is percipient of impermanence (witnessing). In casting our nature as entirely conditioned, the unconditioned *nibbāna* – with its probable contribution to the sense of self – gets entirely ignored. Hume is famous for casting our nature as nothing more than a bundle of thoroughly impermanent perceptions (etc.) and so it is therefore unsurprising to see deep parallels being drawn between Buddhism and Hume on this matter:

> [Hume concludes] that '[persons] are nothing but a bundle or collection of different perceptions which succeed each other with an inconceivable rapidity, and are in perpetual flux and movement.' … Buddhist writers typically make the same point by analysing a person into the 'Five Aggregates' [*khandhās*]. Since a person is nothing more than the sum of these five aggregates, and since soul, in the sense of a permanent unchanging subject of consciousness [viz., the self], cannot be identified with one or more of the five, soul cannot exist … It seems clear that Hume and the Buddhists say the same thing for the same reasons: both analyse the 'soul' [viz., *sense* of self] into a series of events or processes, and do so because this is what experience reveals.
>
> (A.H. Lesser, 1979, 58)

> Moment by moment, new experiences happen and are gone. It is a rapidly shifting stream of momentary mental occurrences. Furthermore, the shiftiness includes the perceiver as much as the perceptions. There is no experiencer, just as Hume noticed, who remains constant to receive experiences, no landing platform for experiences … Suffering arises quite naturally and then grows as the mind seeks to avoid its natural grounding in impermanence and lack of self.
>
> (Varela, Thomson and Rosch, 1991, 60–61)

> Although the Buddha cites various characteristics that something must have if it is to be considered a self, the most important is that of permanence and identity over time. But when we look to our experience,

> there is nothing but impermanence: our bodies, feelings, and thoughts are forever coming and going. In this sense the Buddha is in complete agreement with Hume: where there is diversity there can be no identity.
>
> (James Giles, 1993, 186)

Further examples, while not making explicit comparisons with Hume's 'bundle theory' of persons, reveal this 'Humean' interpretation of Buddhism to be widespread.[17] Among the following scholars, Rāhula, Kalupahana and Gethin are well known; their work often consulted by other scholars (including authors of popular books on Buddhism):

> What we call a 'being', or an 'individual', or 'I', according to Buddhist philosophy, is only a combination of ever-changing physical and mental forces or energies, which may be divided into five groups or aggregates (*pañcakkhandhā*).
>
> (H. Walpola Rāhula, 1996, 20)

> Buddhist thought presents these five [conditioned] aggregates as an exhaustive analysis of the individual. They are the world for any given being – there is nothing else besides.
>
> (Rupert Gethin, 1998, 136)

> The whole human personality, according to Buddhism, is nothing more than the effectively functional psycho-physical organism. The whole endeavour of the Buddha and Buddhism is to make one realise one's own personality and existence in terms of these unenduring and dependently arisen factors ...
>
> (Hari Shankar Prasad, 2000, 139, note 3)

> Therefore, he [the Buddha] undertook the task of redefining the concept of man. According to him, this was merely a 'bundle of perceptions' (*sankharapunja*) or a group of aggregates (*khandhā*), not discrete and discontinuous, but connected and continuous by way of causality, a 'bundle' (*kāyā*) which, for the sake of convenience, is designated by such names as Sariputa and Moggallāna.
>
> (David J. Kalupahana, 1976, 39, note 4)

The following exchange between the monk Nāgasena and the king Menanda is sometimes cited in support of this interpretation:

> … The Venerable Nāgasena replied: 'I am known as Nāgasena, your majesty, and that is what my fellow monks call me. But though my parents may have given me such a name, it is only a generally understood

> term, a practical designation. There is no question of a permanent individual implied in the use of the word.'
>
> King Menander was perplexed. 'How can you declare,' he said, 'that there is no permanent, individual identity implied in the use of the name, 'Nagasena'?'
>
> Nāgasena asked, 'How did you come here – on foot, or in a vehicle?'
>
> 'In a chariot.'
>
> 'Then tell me, what is the chariot? Is the pole the chariot?'
>
> 'No,' replied the King.
>
> 'Or the axle, wheels, frame, reins, yoke, spokes, or whip?'
>
> 'None of these things is the chariot,' the King replied.
>
> 'Taking the separate parts all together and laying them down on the ground, side by side – is this the chariot?'
>
> 'No again,' said the King.
>
> 'Then is the chariot something separate from all its parts?'
>
> 'No, your reverence.'
>
> 'Then for all my asking,' concluded Nāgasena, 'I can find no chariot. The word 'chariot' is a mere sound. You know what the word chariot means. And it is just the same with me. It is on account of the various components of my being that I am known by the practical designation 'Nagasena'.'
>
> (Miln. 25)[18]

The purpose of this passage is to point out the basis upon which one thing is to be distinguished from another, allowing us, in practical contexts, to designate various objects with the use of conventional names. It points out that in the case of persons, we must not be led to believe that any 'permanent individual identity' (viz., a self-entity) lurks behind the use of such proper names as 'Nāgasena' – any more than a chariot-essence lurks behind the use of the term 'chariot'. It is on account of the *khandhā*-components, rather than some permanent individual essence, that we are able to designate, for practical purposes, objects such as chariots and persons by their conventional names. Now we should note that *nibbāna*, being undifferentiated, would play no specific role in the designation of one thing from another, and so there would be no need to allude to it in the context of Miln. 25. Yet if one does not acknowledge the role that *nibbāna is* likely to play in our overall constitution as persons, via the consciousness *khandhā*, then it is easy to interpret Miln. 25 as being about the holus-bolus nature of persons, reading into it the idea that we are nothing *but* a bundle of thoroughly impermanent aggregates. But if we look carefully, Miln. 25 is not saying this at all; it is merely pointing out that the conditioned *khandhās*, as opposed to individual essences, are what allow us to designate persons in a practical context. It is not endorsing a Humean-like 'bundle theory' of persons.

Conclusion

Chapters 2 and 3 offer an interpretation of Theravada Buddhism that relates a particular notion of witness-consciousness – implicit in the suttas – to its well-known principle of no-self (*anattā*). In summary, this kind of consciousness, not to be equated with the consciousness *khandhā* (although an element of it), is intrinsically of an unconditioned, percipient nature that I have termed 'nibbānic consciousness'. It is the underlying 'luminous' nature of mind. The nature of this witnessing *as* unconditioned will be known to those acquainted with its full percipience; *nibbāna* as objectless and subjectless. Lacking the normal structure of consciousness, 'pure' nibbānic consciousness will lie beyond the scope of imagination; yet it will be experienced.

Because *taṇhā* and identification serve to cover and distort the presentation of its unconditioned nature, nibbānic consciousness is only partially revealed through the sense of self. Its presence comes through in the features of witnessing, unity, unbroken presence, longer-term endurance, invariability, unconstructedness and in the ideal of a *dukkhā*-free happiness that is sought after. I have used the term 'pre-nibbānic consciousness' to describe, on the Buddhist position, the witnessing and its various features as partially revealed through the self-sense. Pre-nibbānic consciousness is hence nibbānic consciousness reflexively assumed to be a self – or, to put it another way, nibbānic consciousness as 'covered' or 'duped'.[19] This assumption on the part of pre-nibbānic consciousness will be a mistaken one, one that must be erased if *dukkhā* is to cease. The excision of the self-sense from a person's mindset – such that the mind loses all sense of 'me' or 'mine' towards the *khandhās* – will imply that the intrinsic nibbānic witnessing is percipient of its nature *as* 'uncovered'. Lacking this extrinsic 'covering', it will no longer be *pre*-nibbānic consciousness, but *proximate* nibbānic consciousness (when objects are apprehended) or *pure* nibbānic consciousness (when objects are not apprehended).

Having extracted from Buddhism an interpretation of no-self, the task is now to defend it. In a nutshell I will be arguing that the self is an illusion fed by two tiers – a tier of non-illusory witness-consciousness and a tier of *taṇhā*-driven thoughts and emotions – with the deeper metaphysical status of the witness-consciousness being left as neutral. Whether witnessing is unconditioned or merely a byproduct of the brain should make no immediate difference to my analysis. What is important to recognise, however, is that nothing in the concept of witness-consciousness, as described in Chapter 1, obviously *rules out* the possibility of its being unconditioned. Hence, although grammatical convention can make witnessing sound like a contingent state or activity of a subject, we are not to take this manner of speaking as indicative of any deeper reality. I want this enquiry to *leave the door open to future analysis*, which may well favour witnessing as being unconditioned.

The aim of this book is hence to show that the Buddhist-derived 'doctrine' of no-self – that treats the illusion of self as two-tiered – is a highly plausible account that deserves independent recognition in the philosophy of mind. I will argue that the 'tier' of witness-consciousness is conceptually tied to those features regarded in Buddhism as 'pre-nibbānic'. I then try to demonstrate that this intrinsically unified, unbroken, invariable (etc.) witness-consciousness is real in and of itself, but that, through processes of *taṇhā* and identification (tier two), it mistakenly assumes itself to be part and parcel of a personalised, bounded self-entity. I will thus argue (with some support from neurological research) that Buddhism is right to suppose that *taṇhā* and identification, in their relation to unified, unbroken (etc.) witness-consciousness, play an active role in creating the mistaken impression of a self. The first step in defending this Buddhist doctrine of no-self is to demonstrate that we commonly do, indeed, take ourselves to be an entity that Buddhism alludes to as 'self'.

4
The Reflexively Assumed Self

1. An East–West convergence on the description of self

We have extracted the main notion of self that is implicated in Buddhist suttas. It is a self that Buddhism supposes most people to have a sense of being and, furthermore, a self that as a whole lacks reality on the Buddhist metaphysical system (although a subset of its ascribed features, sourced in nibbānic consciousness, do have grounding in reality). In this chapter, I investigate whether most subjects plausibly do, in fact, reflexively identify themselves as such selves. I hence look to whether we as subjects have a *sense of* being the kind of entity alluded to in Buddhism. If we do indeed harbour this sense of self, as this chapter will argue, then in later chapters we can address the ontological question: are we really the self-entity that we assume we are? If the self is real, then this sense of self will reflect a real self that is sensed. If the self lacks reality, then our sense of self will not reflect an actual self but will have to be explained by factors that do not appeal to the self in their ontology.

To recapitulate, the self alluded to in Buddhism is of the following description:

> A self is defined as a bounded, happiness-seeking/*dukkhā*-avoiding witnessing subject that is a personal owner and controlling agent, and which is unified and unconstructed, with unbroken and invariable presence from one moment to the next, as well as with longer-term endurance and invariability.

How do we go about discovering whether we assume ourselves to be an entity with those specific features? We saw in Chapter 1 that the details of such an assumption, if correct, are unlikely to be introspectively obvious. The purported self conveys the impression of being essentially *subject-like* rather than object-like: it will therefore seem, as a whole, to elude introspective attention. This does not mean that introspection will serve *no*

purpose in determining features of the assumed self. For a sense of self still has a phenomenal dimension (as mentioned in Chapter 1) – it is just that its phenomenal character will seem to lie outside the purview of direct attentional scrutiny. The very elusiveness of this phenomenal dimension will itself figure in the self's description while, at the same time, it will underscore a need for further methods to be recruited when discerning various features ascribed to the self. Such methods will include looking to common patterns of behaviour, linguistic practices, emotions, motivations and philosophical puzzles that may arise.

A further highly useful source of evidence, although this will be *indirect* evidence, will lie in passages from Western philosophers (to be quoted soon) who have sought, for various reasons, to extract notions of the self we assume we are. Suppose that the emerging notion of 'ordinary' self that Western philosophers are concerned with coincides, for the most part, with the notion that concerns Buddhism. This will serve as strong initial (although indirect) evidence that Buddhism is indeed alluding to an assumption of self that is, as they claim, common to humanity rather than something that turns out to be peculiar to their tradition. It will also situate the concerns of this project as being both embedded in the history of Western philosophy and of lively interest to contemporary debate.

Of the philosophers I shall quote, some are explicitly setting out to defend their description of self, while others are just assuming it. Yet others are describing the self as part of an attempt to defend or refute claims about its existence. This is all fine – as long as we keep in mind that such ontological questions are not the current concern. While these philosophers emphasise different aspects of the self, I will suggest that an overall picture of 'self' will emerge that includes and expands upon the description that I gleaned from Buddhist literature. From these quotations, I shall therefore extract a list of key, defining features ascribed to the self. Following this, I will outline each feature, arguing that each is commonly ascribed to that entity which most of us identify as our 'self'. Some of these arguments will be developed further in later chapters. Here then are the selected quotations, beginning, fittingly, with the philosopher who made famous the notion of '*res cogitans*':

> Rene Descartes:
> I am a thinking (conscious) thing, that is, a being who doubts, affirms, denies, knows a few objects, and is ignorant of many, – [who loves, hates], wills, refuses, who imagines likewise, and perceives; for, as I before remarked, although the things which I perceive or imagine are perhaps nothing at all apart from me [and in themselves], I am nevertheless assured that those modes of consciousness which I call perceptions and imaginations, in as far only as they are modes of consciousness, exist in me.
>
> (1641, Meditation III, 42)

… And certainly the idea I have of the human mind in so far as it is a thinking thing, and not extended in length, breadth, and depth, and participating in none of the properties of body, is incomparably more distinct than the idea of any corporeal object …

(1641, Meditation IV, 63)

David Hume:

But self or person is not any one impression, but that to which our several impressions and ideas are supposed to have a reference. If any impression gives rise to the idea of self, that impression must continue invariably the same, through the whole course of our lives; since self is supposed to exist after that manner … There is properly no *simplicity* in [the mind] at one time, nor *identity* in different, whatever natural propension we may have to imagine that simplicity and identity … What then gives us so great a propension to ascribe an identity to these successive perceptions, and to suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives?

(1739, 161–162, 163)

John Locke:

Personal Identity in Change of Substances. That this is so, we have some kind of evidence in our very bodies, all whose particles, whilst vitally united to this same thinking conscious self, so that *we feel* when they are touched, and are affected by, and conscious of good or harm that happens to them, are a part of ourselves, i.e. of our thinking conscious self. Thus, the limbs of his body are to everyone a part of himself; he sympathises and is concerned for them. Cut off a hand, and thereby separate it from that consciousness he had of its heat, cold, and other affections, and it is then no longer a part of that which is himself, any more than the remotest part of matter. Thus, we see the substance whereof personal self consisted at one time may be varied at another, without the change of personal identity; there being no question about the same person, though the limbs which but now were a part of it be cut off …

(1690, 213)

Person, as I take it, is the name for this self … It is a forensic term, appropriating actions and their merit, and so belongs only to intelligent agents, capable of a law, and happiness, and misery. This personality extends itself beyond present existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present. All which is founded in a concern for happiness, the unavoidable concomitant of consciousness; that which is conscious of pleasure and pain desiring that the self that is conscious should be happy.

And therefore whatever past actions it cannot reconcile or appropriate to that present self by consciousness [memory], it can be no more concerned in than if they had never been done …

(1690, 220)

William James:

If the stream [of consciousness] as a whole is identified with the Self far more than any outward thing, a *certain portion of the stream abstracted from the rest* is so identified in an altogether peculiar degree, and is felt by all men as a sort of innermost centre within the circle, of sanctuary within the citadel, constituted by the subjective life as a whole. Compared with this element of the stream, the other parts, even of the subjective life, seem transient external possessions, of which each in turn can be disowned, whilst that which disowns them remains. Now, *what is this self of all the other selves*?

Probably all men would describe it in much the same way up to a certain point. They would call it the *active* element in all consciousness; saying that whatever qualities a man's feelings may possess, or whatever content his thought may include, there is a spiritual something in him which seems to *go out* to meet these qualities and contents, whilst they seem to *come in* to be received by it. It is what welcomes or rejects. It presides over the perception of sensations, and by giving or withholding its assent it influences the movements they tend to arouse. It is the home of interest, – not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak. It is the source of effort and attention, and the place from which appear to emanate the fiats of the will … Being more incessantly there than any other single element of the mental life, the other elements end by seeming to accrete round it and to belong to it. It becomes opposed to them as the permanent is opposed to the changing and inconstant.

(1890, 297–298)

Gilbert Ryle:

When a child, like Kim, having no theoretical commitments or equipment first asks himself 'Who or What am I?', he does not ask it from a desire to know his own surname, age, sex, nationality, or position in the form. He knows all his ordinary personalia. He feels that there is something else in the background for which his 'I' stands, a something which has still to be described after all his ordinary personalia have been listed. He also feels, very vaguely, that whatever it is that his 'I' stands for, it is something very important and quite unique, unique in the sense that neither it, nor anything like it, belongs to anyone else. There *could* only be one of it.

(1966, 31)

Daniel Dennett:
… for ourselves, it seems, consciousness is precisely what distinguishes us from mere 'automata'. Mere bodily 'reflexes' are 'automatic' and mechanical; they may involve circuits in the brain, but do not require any intervention by the conscious mind. It is very natural to think of our own bodies as mere hand puppets of sorts that 'we' control 'from inside' … There are notorious problems with this idea, but that does not prevent it from seeming somehow right: unless there is a conscious mind behind the deed, there is no *real* agent in charge. When we think of our minds this way, we seem to discover the 'inner me,' the 'real me.' This real me is not my brain; it is what *owns* my brain ('the self and *its* brain').

(1991, 32)

Richard Baron:
The self which this paper rejects is not the individual organism … It is the self as one who thinks and acts … there is not really a knowing self, observing the world and learning. There is only an organism with its brain being etched by its encounters with the rest of the world. (2000, sec 2.0) … Denying the reality of self means … not accepting that there is anything special about a particular individual: someone else with the same properties would be just as good … If you look at a person and see only the organism, characterised by its properties, you will not notice if another qualitatively identical person is substituted … there would be nothing special about the individual as such. (2000, sec 4.0) … We must distinguish between the feeling that one is a self and the fact (or rather fiction) that one is a self. We can feel many things to be the case even though they are not the case. (2000, sec 4.1) … the sense of self is taken by the person feeling it to prove that there is a real, enduring, singular self … (2000, sec 6.1)[1]

John Canfield:
What is our very idea or conception of the 'I'? (1990, 19) … Something that is not the body but owns it; something that perceives, thinks, and wills; something that persists over time. This is the core conception of self virtually everyone has or presupposes. By saying we presuppose it, I mean that our way of talking about ourselves and our beliefs about ourselves imply this view; or at least, we would normally take it to be true, on reflection, that there is such an implication.

(1990, 20)

Galen Strawson:
What, then, is the ordinary, human sense of self, in so far as we can generalise about it? (1997, 407) … I propose that the mental self is ordinarily conceived or experienced as: (1) a *thing*, in some robust sense (2) a *mental*

thing, in some sense (3,4) a *single thing* that is single both *synchronically* considered and *diachronically* considered] (5) *ontically distinct* from all other things (6) a *subject of experience*, a conscious feeler and thinker (7) an *agent* (8) a thing that has a certain character or *personality* (1997, 407–408).

Peter Strawson:
… we have not only the question: Why are one's states of consciousness ascribed to anything at all? We have also the question: Why are they ascribed to the very same thing as certain corporeal characteristics, a certain physical situation, etc? It is not to be supposed that the answers to these questions will be independent of one other.

(1959, 89–90)

Owen Flanagan:
… an 'I' that stands behind all conscious experience [and] … constitutes the core of the self, our conscious control centre, the source of all action plans, and the agent for whom all experiences accrue before being filed for future reference or discarded … (1992, 177) … the mind's 'I' is an illusion. … The illusion is that there are two things: one side, a self, an ego, an 'I,' that organizes experience, originates action, and accounts for our unchanging identity as persons and, on the other side, the stream of experience (1992, 178).

Jonathan Shear:
In the West, philosophers such as Descartes, Hume, Kant and Sartre have been utterly unable to come to anything even approximating agreement as to whether there is, or even possibly could be, anything in our experience which could fulfil our commonsensical conception of self as the unitary conscious subject of one's thought and experience.

(1996, 359)

Antonio Damasio:
Besides those images [of what you perceive externally] there is also this other presence that signifies you, as observer of the things imaged, owner of the things imaged, potential actor on the things imaged. There is a presence of you in a particular relationship with some object. … The presence is quiet and subtle, and something little more than a 'hint half guessed,' a 'gift half understood,' to borrow words from T.S. Eliot … (1999, 10). In all the kinds of self we can consider one notion always commands centre stage: the notion of a bounded, single individual that changes ever so gently across time, but, somehow, seems to stay the same … (1999, 134). Whether we like it or not, the human mind is constantly being split, like a house divided, between the part that stands for the known and the part that stands for the knower (1999, 191).

It should be evident from these passages that a split between a knowing and conscious *subject* and the known *object* (as I have defined the terms) is absolutely basic to this notion of ordinary self. Metaphorically speaking, the subject might be thought of as the grain of sand upon which the self-pearl is grown. The subject becomes characterised as 'the self' through a collection of roles (e.g., 'observer', 'owner', 'actor') and attributes (e.g., 'conscious', 'unified') that 'bind' themselves to the subject. Some of these roles and attributes seem intrinsic to subjectivity itself; others, arguably, are additional. I would like to suggest, therefore, there is something like a common idea of the self that can be seen in the above writers. In the several sections that follow, I will try to outline this idea with a view to show that this common idea describes the self that we assume we are.

1.1. Roles ascribed reflexively to the self

In this common idea of the self, we reflexively take the self to be a *subject* that is a:

- Knower/observer/witness of experience, source of attending
- Owner of thoughts, perceptions, experiences, body, personality
- Initiator or agent of actions – the source of effort and will
- Thinker or originator of thoughts
- Seeker of happiness over suffering

1.2. Attributes ascribed reflexively to the self

In the common idea of the self, we also reflexively take the self to be a *subject* that is:

- Conscious/mental/aware for much of the time
- Bounded, viz., ontologically unique in its identity
- Elusive to attention ('hint half guessed'), vague
- Unified/singular/simple
- Uninterrupted and unchanging in its essence – both from one conscious moment to the next and over longer tracts of time (responsible for our identity)
- Unconstructed by – and standing apart from – observable objects in the 'stream' of consciousness

Each of the phrases for these roles and attributes will receive explanation and amplification in the following sections.

Readers will note the obvious similarity to emerge between the notion of self alluded to by Western philosophers and that alluded to in Buddhist suttas. Features such as 'thinker of thoughts' and 'elusive to observation' are apparent additions, although minimal reflection will show these aspects to fall naturally out of the Buddhist notion. Hence, anything that is subject-like

rather than object-like will elude its own attentive observation, and the role of 'thinker of thoughts' will be implicated in the closely related role of 'agent' or 'initiator of actions'.

With this background, I will now provide more direct arguments for the claim that most people reflexively assume themselves to be an entity with such roles and attributes. Where necessary, I elaborate on what these features involve, or point to further distinctions such as that between perspectival and personal ownership (both of which are ascribed to the self). The roles and attributes are not to be considered independently of each other, since some of them seem to be inextricably linked – for example, the roles 'knower/observer/witness' and the attributes 'conscious/mental/aware'. Where linked, I consider them in relation to each other, under the rubric of a general section. Importantly, it should be borne in mind that in arguing that most people assume themselves to be an entity with each of the above-mentioned facets, I am *not* arguing or supposing that any such facet has independent reality. And when discussing each of its supposed facets, I will use the term 'self', unless otherwise specified, in a *general* way to refer simply to *the thing we assume we are*. This will avoid prematurely supposing that we take ourselves to be that entity with all the aforementioned features. Whether the thing we assume we are amounts to (say) a mere impersonal subject – or to more than just this – will be determined as the discussion unfolds.

2. Role: Knower/observer/witness; Attribute: Mental/aware/conscious

The role 'knower/observer/witness' would appear to best characterise that which makes the subject of experience *subjective* and hence something with a mental or conscious inner life. It was noted in Chapter 1 that the *modus operandi* of the subject, even if not active all the time (such as during deep sleep), is that of *witnessing*, this being the broadest mode of conscious apprehending, attentive and inattentive. What helps contribute to the impression that the witnessing subject stands apart from objects (and is of the class, 'individual') is the fact that, while presented with objects, witnessing appears to emanate from a particular psycho-physical perspective of a body and mind. The subject is the locus of our first-person perspective on the world; it is witnessing as it presents from such a perspective.

From its outset, this project has supposed that a subject/object distinction, as outlined by Mait Edey (1997, online) and elaborated in Chapter 1, is implicitly taken on board by anyone with an inner conscious life who interacts with the world and its objects. A section of the passage from Edey states:

> You can distinguish yourself as subject from any object whatsoever ('physical' or 'mental') any time you direct your attention to that object and realize that it is you who are aware and who pay attention, not the

> object. The real nature of the object and the real nature of the subject may be baffling mysteries to us, but these mysteries are no barrier whatsoever to knowing which is obviously which.
>
> (1997, online)

As this passage suggests, the fact that most of us pre-theoretically distinguish ourselves as a conscious, witnessing subject in opposition to observed internal or external objects seems obvious enough, and is indeed taken for granted by many authors of the above quotations. It is an assumption that underpins the way we use language. When we use the first-personal pronoun 'I', for instance, there are many contexts in which this 'I' seems not to refer to any *object* in our purview, but to ourselves as the *subject* that observes those objects.

If pushed for further argument that we commonly make this distinction, I would suggest that the very notion of first-person memory involves an appeal to a witnessing subject. (I am not presuming here that the notion of 'first-person memory' picks out anything real in the world – rather I am appealing to what seems, ordinarily, to be the case.) When we recall experiencing an event, for instance, our memory is not considered genuine unless we really did experience that event. In order to have experienced the event, that event will have had to have stood in what John Perry (1997) would call an 'agent-relative role' by appearing, in either a direct or mediated fashion, to a standpoint that is relative to a conscious subject in a given space and time. For the event to appear this way, such that it can be consciously recalled later, it will have to be consciously (as opposed to merely unconsciously or subconsciously) apprehended at the time. As I see it, the only way to make sense of 'consciously apprehend an event' – in such a way that the event can be explicitly recalled later – is to construe the apprehending as involving a form of *conscious mental recording*: as involving, in other words, conscious observing or witnessing. That we can consciously recall an experience seems to crucially depend upon the fact that we, as conscious subjects, witnessed it to begin with. It notably does not seem to matter if the recalled experience pertained to an object that was, at the time, seen, heard, tasted, smelt, felt, imagined or remembered. What matters to this argument is not the specific perceptual or introspective mode through which the objects are witnessed, but the fact that they were consciously witnessed, whatever the mode. Such reflections reinforce the contention that there seems to be a common or *generic* aspect to such modes of witnessing, in virtue of which they are *able* to be recalled. Further evidence for the generic nature of witnessing was cited in Chapter 1. The very fact that we are able to know, immediately from the first-person perspective, *that* we are seeing, hearing, feeling, or introspecting thoughts is suggestive of a type of conscious generic knowing that is involved in but not equated to any specific mode (e.g., we do not *hear* that we hear; we *know* that we hear).

The word 'conscious' (as opposed to 'unconscious' or 'subconscious') has been mentioned several times. I use the term 'conscious experience' to convey something phenomenal, such that there is, in Thomas Nagel's (1986) phrase, *something it is like* to have that experience – or, perhaps, to be a thing which has that experience. However we construe it, an inner phenomenal life seems central to what it means to be a subject whose *modus operandi* is witnessing. That is, the subject of the above characterisation, which seems at the core of the self, is not merely a lifeless thing, like a rock, but is predominantly a mental or conscious thing with a subjective inner life. Regarding this subjective inner life, we might further ask: is there something it is like to be a subject or self *per se*, such that the witnessing *qua* self seems imbued with its own intrinsic feel, and whose character is a part of any given experience? Or does the phenomenal character of a given conscious experience seem exhausted by the qualities of the (attendable) *objects* that are witnessed, including the characters of different perceptual or introspective *modes* such as that of seeing or imagining?

Reflection upon the nature of conscious experience suggests the former to be true. We naturally forge – and allude to – a subject/object distinction (or as this chapter will suggest, a *self*-other distinction) on the basis of *first-person experience*, that is to say, on the basis of how things seem from our own point of view. Why would we allude to such a distinction if our experience seemed exhausted by its objects? It will transpire during this chapter that inherent to the harbouring of a sense of self will be an implicit impression that *the self* (namely, oneself) is a subject-like, bounded entity that seems to stand apart from its objects. It will partake, in what we noted in Chapter 1, in an elusiveness to its own attentive purview. And yet this subject-like entity, the self, will not seem like a mere abstract inference, but will seem like something whose existence and character is gleaned, if only elusively and indirectly, through first-person experience, such that there seems *something it is like* to be a subject *qua* self. This is not to make the claim that this supposed phenomenal character possessed exclusively by the self is *in fact* intrinsic to a subject or self (as opposed to objects of experience); the claim is only that it *seems* to be the case.

3. Roles: Owner, agent, thinker, seeker of happiness; Attribute: Bounded

The Buddhist analysis of self in Chapter 3 cast as central the role *personal owner*. If we naturally assume what I have termed a *self-identity* with this role, we as subjects will implicitly take ourselves to be a self with a *boundary*. By 'boundedness', I mean that the entity in question is ontologically distinct from all other things. It is not an impersonal perspective from which objects are witnessed (as implied by the bare notion of 'subject') but a personalised thing with numerical identity. While the self's boundedness

(where 'the self' ends and 'the other' begins) will not be directly observable to the subject who identifies as a *personal owner*, the assumption of boundedness may nevertheless be *evidenced* through the fact that we take some *khandhās* to be 'me', and others to be 'not me'. Boundedness thus underscores a self/other distinction: the casting of self as an ontologically distinct entity. As Ryle puts it:

> He also feels, very vaguely, that whatever it is that his 'I' stands for, it is something very important and quite unique, unique in the sense that neither it, nor anything like it, belongs to anyone else. There *could* only be one of it.
>
> (1966, 31)

Boundedness is hence that feature by which the subject *qua* self has a unique identity that separates it from all other things. To have a *sense* of being a bounded entity is to have a sense of *being* a unique thing with boundaries separate from all other things.

In this chapter, since boundedness is such a central and important feature of the self in Buddhism, some time is spent arguing that most of us do indeed identify as a bounded self-entity, and that the role *personal owner* is central to the assumed boundedness of self. During the course of this discussion, the roles *thinker/agent* and the attribute *seeker of happiness over suffering* are brought into play, with the conjectured relation between boundedness and *taṇhā* discussed at the end. This discussion of boundedness proceeds in the four sections (Sections 3.1–3.4) as follows.

Section 3.1 begins by saying more about the notion of identification. Its main goal is to show, in general terms, how identification serves as the best evidence for the fact that we take ourselves to be bounded entities. While this draws upon the Buddhist analysis of Chapter 3, I no longer use Buddhist terminology (e.g., 'the *khandhās*'); nor do I touch on all the same details (which supplement this current analysis). Section 3.2 outlines four common modes of assumed self-identity, which are also powerful pods of evidence for our ascription of boundedness to the self. These modes are 'this-ness', 'thinker/agent', 'consistent self-concern' and 'personal owner'. In these modes, the subject identifies itself *with* various items (and hence *as* a self) *in the capacity of* the modes. The assumed self-identity roles 'thinker of thoughts' and 'agent of actions' are demonstrated in this exercise, with the attribute 'seeker of happiness over suffering' addressed under the rubric of 'consistent self-concern'.

The fourth mode of *personal owner*, I argue, is the broadest; I therefore spend some time on this. So the sutta in which personal ownership (or 'myness') is co-defined with a self (or 'me') is revisited. I argue, in accordance with the sutta, that 'where there is a (sense of) mine, there is a (sense of) me' and 'where there is a (sense of) me, there is a (sense of) mine'. In the course

of this exercise, I argue that the first three modes of assumed self-identity – this-ness, thinker/agent and consistent self-concern – all imply a sense of personal ownership (or my-ness) and hence, in turn, a sense of being a personal owner. The role *personal owner* is therefore demonstrated, in accordance with Buddhist analysis, to be the most ubiquitous mode of assumed self-identity – central to the assumed self – as well as the best role to evidence a sense of boundedness.

Having discussed the main modes of assumed self-identity (this-ness, agency, consistent self-concern and personal ownership) whereby a subject identifies with various items in the capacity of such modes, I turn, in Section 3.3, to how these modes and items might become integrated with a subject's overall self-conception (as a personality). I suggest that each interaction between subject and world will reactivate one's general sense of self-identity, such that one does not forget 'who one is'. Finally, in Section 3.4, I consider preliminary evidence for the Buddhist claim that the reciprocal senses of self-identity and personal ownership co-arise with *taṇhā* (emotional investment in desire satisfaction) so that whenever one phenomenon is present in the mind, then so is the other. This raises the question – to be pursued earnestly in Chapter 8 – of whether (as Buddhism would hold) the sense of the self's boundedness could, contrary to how it seems, actually *amount to* (as opposed to just be evidenced by) a subject's identification with various items, through the assumed modes of self-identity.

A note on the term 'subject': appeal to this term is unavoidable when explaining such notions of identification. Readers should not take this to mean a subject has been established to have independent reality: only that identification is to be analysed in terms that appeal to a subject. While Buddhism takes such a subject to be real, should the subject turn out to lack reality, then so, too, will identification, etc.

3.1. Identification as general evidence for boundedness

> 'Arnie identifies with his fit body'.
> 'How are you Julie?' 'We're doing great. We're pregnant!'
> 'The blue hair and body-piercings are part of who I am. I feel incomplete without them'.

Identification occurs when a subject 'appropriates' various things to its perspective, so that they do not overtly *appear* as things separate to that subject (namely, the witnessing-from-a-perspective). From the subject's point of view, it is as if the identified-with items become assimilated to the subject (this fit body, this blue hair), or the subject becomes part of a larger item (that doting couple where the pronoun 'I' entirely disappears from conversation with the outside world). And when the union of subject and appropriated item, identified with as 'me' or 'we', confronts the world as 'self' or 'us', this world is

invariably thrown into relief as 'other'. This sense of separation between 'self' and 'other' is evidence for a feeling of boundedness that is had by the identifying subject. Identified-with items may include gender, race, character traits, preferences, aversions, bodily attributes, social status, profession, possessions, other people, political affiliation, sports teams and countries. It can be as simple as a perspectival owner of one's perception or as complex as a Marxist manifesto. In cases where a subject identifies itself *with* or *as* something so overtly non-mental as a football team, it will really be only the *thought* or *idea* of that thing that could get appropriated, in any literal sense, to the subject's perspective. In other cases, a subject may appropriate to its perspective various ideas *about* its own mental life (e.g., that of being a hot-tempered person), or it may appropriate more immediate (less representational) mental phenomena, such as joy, pain or those psycho-physical cues that lend witnessing its perspective on the world. I will henceforth use the term 'appropriates' (along with 'identifies') to convey what a subject actually *does* upon identifying itself with or as an item. Whether any such ideas or mental phenomena are in fact *assimilated* with the perspective of a subject is a matter of contention.

Given that there is witnessing, identification seems to subtly affect the manner in which things are witnessed, such that one's thought and perception are imbued with the distinctive flavour of 'oneself'. J. David Velleman (2002) describes identification along these lines, and although he limits his discussion to aspects of the personality, his analysis also works with aspects such as physical attributes (or our idea of them). In his paper 'Identification and Identity' (2002) he suggests that when there is a part of one's personality that 'presents a reflexive aspect to [one's] thought', that part of the personality is identified with as 'me', or is, as he puts it, as part of one's self-conception at that time (Velleman, 2002, 114). To present a reflexive aspect to one's thought, according to Velleman, is to view the world through the lens of that aspect, such that its filter seems to become a part of the subject's first-person perspective. In his own words:

> If there is a part of your personality with which you necessarily think about things, then it will be your mental standpoint, always presenting a reflexive aspect to your thought. You will be able to think about this part of your personality as 'it,' but only from a perspective in which it continues to function as the thinking 'I' – just as you can find a reflection of your visual location 'over there' only from a perspective in which it is also 'back here.'
>
> (2002, 114)

So if a subject (of which we shall call the associated *person* 'Arnie') identifies itself as a physically fit person, then so long as this identity persists, Arnie cannot help but think about and perceive the world as a *fit person*. Arnie's perspective on the world will be imbued with this idea in such a way that it shapes his pattern of behaviour and desires (e.g., visiting the gym regularly).

It will shape the way he thinks, including his emotions – for example, when hearing the statistics of how many morbidly obese adults there are, a surge of pride may accompany the thought '*I* am not one of those people'. And as Velleman rightly points out, so long as an aspect of the personality is identified with as 'me', one will not be able to think about that aspect from a truly detached perspective. Arnie will always think about his fitness through the filter of his conception *as* a fit person. Identification is essentially reflexive in this manner. In many cases a subject will not be overtly aware that it is identifying as this or that thing. Yet the identification, such as with one's gender, or as someone with free-will, will be 'colouring the lens', modifying the way the world is perceived.

Identification seems to be something that we do habitually. I propose that these identity filters (of which there will be many at a time) would, in broad terms, evidence a felt boundary between what seems to be our *self* and the world in the following way. When a subject confronts the world, including the world of bodies, thoughts and perceptions, a boundary that is felt to wordlessly differentiate this subject (viz., witnessing-from-a-perspective) from the witnessed objects is also, or rather *actually*, felt to differentiate the subject *qua* me (via appropriated ideas), from encountered objects not identified with as 'me'. The subject who confronts an object as 'other', then, does not normally confront the object in the passive manner of an impersonal witness that happens to take a psycho-physical perspective on the world. It confronts those objects actively, *as* those features that seem to adorn its perspective: *as* the body, mind, personality, favourite sports teams, and so forth. Dennett (1991, 417) has noted that in a certain respect, the boundaries of 'self' do not appear static, but as fluid, expanding and contracting with context. In one situation, one's sense of identity may expand to include a favourite sports team, feeling that one personally has a stake in winning or losing the match ('the other', of course, being the rival team and sometimes, country). In such situations, one's identification with an item may be described in terms of being 'part of that item' with the words 'we' and 'our' being common, as in 'we won the match'. In another situation, one's sense of identity may shrink away from one's thoughts and actions to the point of dissociation, as in 'That wasn't the real me talking' (Dennett, 1991, 417). Whether expanded or shrunken, the sense of boundedness, through identification, seems to turn the subject from a mere point of view into a *substantial personal thing*.

3.2. Four common modes of assumed self-identity: This-ness, autonomy, consistent self-concern and personal ownership

Now that we have some idea of how a sense of boundedness is evidenced through identification in general, I present further evidence that makes clearer the fact that we as subjects regard ourselves to be bounded and distinct entities. This evidence points to repetitive *modes* or *ways* in which we

as subjects identify with various items, and hence as a *self* in their capacity ('the self' referring to the subject as assimilated with items of identification). These modes of identification will offer good reason to suppose that we assume there to be a thick psychological boundary between self and other. Our reflexive ascription of ontological uniqueness (or boundedness) to ourselves – as something distinct from all other things – is evidenced through the following modes of identification in which a *self* seems presupposed. There is: a unique importance or value felt in association with simply being, it would seem, *this very thing* (Section 3.2.1); a felt *causal efficacy* (through seeming to be a free-willed initiator of actions, thinker of thoughts) (Section 3.2.2); a pattern of thoughts, emotions and behaviour that seem, over time, to reflect a concern for one's own self (Section 3.2.3); and the sense of being a personal owner-self (Section 3.2.4).

3.2.1. The importance of being 'this very thing'

While Galen Strawson (1997) lists 'ontological independence' as part of the ordinary conception of self, he does not detail how this is to be construed pre-theoretically, as a feature most of us take for granted. Both Ryle (1966) and Baron (2000), however, make more explicit reference to this, which Baron refers to as 'this-ness': a feature to which we shall now turn.

Baron (2000) alludes to the self (whose reality he rejects) as something that essentially has a quality of specialness or, as he calls it, 'this-ness' (2000, sec 4.2.3, 5.0 & 7.0). 'This-ness' is conveyed through the feeling that substituting oneself with another individual who has all the same qualities leaves out something essential – the value associated with being 'this very thing' which is 'me'. 'This-ness' will involve identification on the part of the subject, sometimes with a specific object, like an author, but always as accompanied by a more general idea: that of *being the one and only me* who is the author, etc. This-ness hence reflects a mode of self-identity that can be commonly assumed by a subject. In all such cases, the subject's identification *as* the one and only me is suffused with a dispositional sense of *value* that can be readily felt, should the occasion arise. This value seems to create or reflect an almost tangible feeling of boundedness between self and other. An example Baron (2000, 7.1.1) uses to illustrate such feeling involves that of personal achievement: the desire that something, such as a project, be done by oneself and oneself alone. For someone writing a novel, for instance, it is generally of no comfort for them to learn that someone else, no matter how similar in attributes, is to take over its writing. It matters to him that he, as opposed to someone else, writes the novel. He identifies as its author, and part and parcel of this identification is that he believes and implicitly desires it to be *he and he alone* who is the author, as opposed to some other subject of experience, no matter how similar. Baron (2000, 7.1.1) invites us to contrast this with other tasks, such as applying the brakes of a car, by which we feel it does not particularly matter *who* does it (although

I contend that if there was an accident, it *would* matter who did it, showing the value dispositionally attached to this action).

The sentiment of 'this-ness', involving identification as 'the one and only me' (and hence a sense of self-identity with this role) is not confined to authoring actions, but includes the value that we ascribe to being *this very subject* of experience. This dimension is sharply conveyed in a thought experiment by Derek Parfit (1984, 199–201). Imagine that a new form of transportation is invented, by which a machine scans every detail of your psycho-physical constitution, replicating your person at the desired destination – before destroying the original. One day, the machine malfunctions, leaving the 'original you' (which may itself be a replica) undestroyed. You talk to your doppelganger over a video link, who informs you that the procedure damaged your heart, leaving you with just a week to live. He tells you not to worry however, since he will be living your life, finishing your book, caring for your wife and children! Any feeling of discomfort you are bound to have is telling of the fact that you value 'you' (and for that matter, others) not only for your (or their) qualities, which are present *holus-bolus* in the replica, but for simply being, it would seem, *that very thing* (even if *that very thing* is itself a replica from previous travels). One identifies, in other words, as 'this *very* subject of experience', as the *bearer* of one's psycho-physical qualities. One can feel this sentiment to the extent that the thought of losing one's 'this-ness', such as through replication and destruction, feels akin to the thought of dying. It shows the great deal of importance one places, whether rationally or not, on the idea of numeric distinctness – over and above any qualitative features. Reflecting on the situation through filters heavily associated with 'being me', the average subject is thus not thinking: 'None of that is who I am, so what does it matter if a replica takes over?' No. He is thinking 'I am the subject of these qualities. If a replica takes over, he lives and I die'. One's ascription of value to being 'this very thing' is hence strong evidence that most people deeply identify as being 'the one and only me' and hence, as an independent subject of experience. It provides strong evidence for a sense of wordless boundary that is drawn between what seems to be the self and all other things, no matter how similar in quality.

3.2.2. Role: Agent of actions and thinker of thoughts (autonomy)

The sense of boundedness is also brought out through considerations pertaining to 'the self's' causal efficacy. For most people take themselves to be autonomous agents in virtue of their assumed causal powers, thus relating directly to the 'role-occupiers' *thinker of thoughts*, *initiator of actions*. These roles point to common modes of assumed self-identity. How, more precisely, do we identify as such thinkers and agents? One way, already mentioned, is through 'this-ness': the felt value attached to the idea that *I*, *this particular self*, as opposed to some other self, am the

agent of certain actions. Another way we construe ourselves to be thinking agents is through the feeling that our deliberate actions are not the result of impersonal factors but, rather, of special causal powers pertaining to free-will – *our* free-will. We feel, in other words, that our choices are not blindly determined, but that with any deliberate action, we could have chosen to do otherwise. The feeling that one is able to exert unique causal powers on the world through one's own thoughts and actions adds weight to the feeling of *being* a separate, autonomous entity. Identifying as a (free) thinker and agent would thus plausibly evoke a sense of boundary between our 'free' selves and the world with which we interact (including other free agents).

But the feelings of freedom do not seem to stop there. Like 'this-ness', the belief in one's free-will seems to endow those free thoughts and actions with value. One takes particular pride or shame not only in the apparent fact that *this*, as opposed to some *other* self, is the rightful author of an action, but in the 'fact' that as an island of special causal power, one is able to author *this*, as opposed to *that* kind of action. It is through this feeling of freedom that one feels responsible for one's actions. In the extensive literature on free-will, it has been noted that anyone who *truly* believed there was no real choice in the matter – that our every action was determined from birth – would not fully experience the emotions of pride, shame, guilt, praise or blame, to name but a few. It seems that for these emotions to be properly felt, one must, at *some* level, buy into the assumption that it is possible to have chosen differently. We do not usually attribute heartfelt praise or blame to behaviours we perceive as mechanistic or random (if we do, then it tends to be through unconsciously anthropomorphising inanimate objects such as stalling cars and red traffic lights!). The emotional investment in the outcome of one's actions serves to intensify the sense of boundary between self-as-agent and other (or self-as-thinker and other). The associated roles, 'thinker of thoughts' and 'initiator of actions' thus depict distinct and repetitive *modes* in which we, as subjects, identify with things (in the capacity of these roles), underscoring the sense of boundary between self and other. And the associated sense of boundedness is best evidenced through the value we attach to being, it would seem, a free author of our actions.

The reflections developed in this discussion on both 'this-ness' and 'autonomy' (introduced by Baron) help to illuminate, from two different angles, the sense of ontological uniqueness that we have. That sense of being a uniquely separate *thing*, whether as something special, or as something autonomous, is strong evidence for our reflexive ascription of boundedness to the self we assume we are. We can also note its connection with the long-running debate on free-will, and with the fact that many philosophers, such as Kant and Frankfurt, have chosen to identify the most central aspect of our 'selves' with 'the will'.

3.2.3. Consistent self-concern

> We know how little it matters to us whether *some* man, a man taken at large and in the abstract, proves a failure or succeeds in life, – he may be hanged for aught we care, – but we know the utter momentousness and terribleness of the alternative when the man is the one whose name we ourselves bear. *I* must not be a failure, is the very loudest of the voices that clamour in each of our breasts: let fail who may, *I* at least must succeed. Now the first conclusion which these facts suggest is that each of us is animated by a *direct feeling of regard for his own pure principle of individual existence*, whatever that may be, taken merely as such'.
>
> (James, 1890, 318)

This passage points to a further powerful source of evidence for the fact that we take ourselves to be a bounded thing. The entity that we seem to care most about is that which answers to our 'self'. Such an observation is borne out over time. For we are most *consistently* concerned with the welfare (including the goals and aspirations) of what would appear to be our *self*, that very entity we feel to be 'me'. The claim is not that we are concerned with the welfare of *only* our 'self' (which would imply that we are selfish) – but rather that, over time, our concerns most consistently and potently involve the welfare of what would seem to be our self. This concern is reflected through our emotions, in the way we think and in the courses of action that we choose. The general thought is this. If there is evidence that over time, our thoughts, emotions or behaviour most consistently involve the welfare of what would appear to be one's 'self', then this also counts as evidence that we most value our 'self'. Hence, we *ipso facto* take ourselves to be that ontologically unique entity that is valued in this way.

Let us consider a range of the emotions; for example, joy, sadness, fear, surprise, anger, moral indignation, aversion, lust, anxiety, frustration, guilt, pride, vanity, embarrassment, envy, nervousness, excitement, jealousy, loneliness and disappointment. On the face of it, in order to feel most, if not all of these emotions, one has to identify as an 'I' who is minimally a *subject* of the emotions. One cannot, it would seem, feel guilty without buying into an *I* who feels guilty, or angry, without assuming (if implicitly) that *I* am angry, and so forth. Most, if not all of these common emotions, would thus seem to include as a part of their 'aboutness' not only the object to which the emotion is directed, but an implicit subject who is experiencing that emotion. But more than this, they often seem to exist because the subject desires what is best for *itself*, either in terms of avoiding harm or accruing benefits. Imagine that one could not care less about one's fate. Could one feel guilty or fearful? To feel guilty usually implies a wishing that one had not done a particular deed, which implies, in turn, a preference to have acted in a way that avoided causing one's own (amongst any other) suffering. To feel fearful generally implies that one wishes that a particular

scenario would not come to pass, which again betrays the desire for an outcome that is relatively unharmful to one's own self. Most of the above-listed emotions involve some investment of one's happiness in a preferred state of affairs. Whose preferences? George W. Bush's? Only if it concurs with what one's own *self* would prefer. These emotions are ubiquitous, felt in various grades of intensity many times a day, from desiring to get rid of the smallest itch, to the more weighty desire to rid oneself of a long battle with illness.

The nature and ubiquity of these common emotions, hence, serves as evidence that we assume ourselves to be not a mere subject, but a bounded subject, namely the thing that, over time, we care most consistently about. We might call this mode of identification, best evidenced by the emotions, 'consistent self-concern'.

3.2.3.1. Evidence for 'seeker of happiness over suffering' as an ascribed feature of the self. The evidence for 'consistent self-concern' also lends considerable support to the Buddhist contention that the very nature of the self that we identify ourselves as being is such that if given the choice, it would seek on the whole to avoid suffering and be as happy as possible, where the notion of happiness can be construed broadly to cover a range of positive emotions characterised by a pleasant affective tone, such as joy, the contentment of goal achievement, peace and fulfilment. We leave open such questions as to what kind of lifestyle might be most conducive to pleasant emotion, or which pleasant emotions are the ones most worth having. I think it is uncontentious to suppose that most people would on the whole seek to minimise their own suffering and maximise their own happiness – rather than to maximise their own suffering or simply not care one way or the other. We are not saying that people are not prepared to undergo privations and suffering to achieve their goals, or that such pain and suffering might not be considered to make the goal even more valuable. But note that the goal, if freely chosen, will rarely be one that we think will make us on the whole less happy. An overall preference to seek happiness and avoid suffering, therefore, seems central to the self that we assume we are. It is worth re-iterating Locke's passage which picks up on this aspect:

> This personality extends itself beyond present existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present. All which is founded in a *concern for happiness*, the unavoidable concomitant of consciousness; that which is conscious of pleasure and pain desiring that the self that is conscious should be happy. And therefore whatever past actions it cannot reconcile or appropriate to that present self by consciousness, it can be no more concerned in than if they had never been done …
>
> (1690, 220) [my italics]

Thus Locke alleges that a concern for happiness is part of the nature of the conscious self and that this fundamental concern is partly what motivates one to identify as the author of one's actions. Interestingly, he links the ownership of actions to this concern for happiness. It would seem that on Locke's account, if one did not show any concern for an action, then this would be a sign that he did not regard the action as his own. This accords well with what Buddhism says about personal ownership, with its link to *taṇhā*. We now turn to ownership.

3.2.4. Ownership and its relation to identification

> Bhikkhus, there being a self, would there be for me what belongs to a self?'
> 'Yes, venerable sir.'
> 'Or there being what belongs to a self, would there be for me a self?'
> 'Yes, venerable sir.'
>
> (MN 22)[2]

In Chapter 3, I argued that the Buddhist analysis of self in terms of personal ownership (in the 'belongs to' relation) is one that makes the reciprocal role of *personal owner* the broadest mode of self-identity and hence, central to their concept of a self. I will now argue that the Buddhist analysis (encapsulated in the sutta above) is correct insofar as it fits with the way we think and speak, making good sense of how different kinds of ownership (perspectival, possessive, personal) are to be distinguished. The role of personal owner will indeed be the most central to the assumed self, such that it is implied in the other common modes of self-identity: this-ness, autonomy and consistent self-concern. I approach this exercise through examining each half of the above sutta, summarised as 'where there is (a sense of) mine, there is (a sense of) me' and 'where there is (a sense of) me there is (a sense of) mine'.

Let us review the definitions. There can be 'possessive ownership', dictated by social convention as in 'I own this computer'. There can be 'perspectival ownership' describing the manner in which objects (such as toothaches, sensory qualities and thoughts) appear to a subject's perspective in a direct way that they appear to no-one else. Some other objects, although publically observable (such as bodies and their actions), are presented to the subject in a unique manner that allows them to be described, more broadly, as being 'perspectivally owned'. And there can be 'personal ownership' (or a sense thereof) whereby various objects – usually perspectivally and/or possessively owned – are felt by the subject to *belong* to a distinct 'me' in a manner that is neither *merely* perspectival nor possessive. A sense of personal my-ness is determined (or reflected) by the subject's *identifying with* the reciprocal roles of perspectival or possessive owner such that it feels at one with them. In this capacity the subject identifies *as* a personal owner of the items in question. The analysis receives amplification by way of examples below.

3.2.4.1. Where there is 'mine', there is 'me' (viz., the sense of being a personal owner). In this sub-section I suggest that a *sense of* personal my-ness – the kind that is lost in episodes of anosognosia or depersonalisation – is most usefully analysed with reference to a reciprocal *sense of* self-identity as a me *qua* personal owner. This sense of self-identity involves identification, on part of the subject, with the roles 'perspectival' or 'possessive' owner.

In Chapter 3, I talked about how a sense of personal ownership does not seem to reduce merely to the perspectival or possessive relation that a subject may bear to objects in the world. Nor does a sense of self-identity seem to reduce merely to a subject's occupation of the reciprocal roles 'perspectival owner' or 'possessive owner'. This is best illustrated through imagining scenarios where the sense of perspectival or possessive ownership comes apart from a sense of personal ownership. Without resorting to anything as extreme as anosognosia, suppose that a self-confessed introvert, Jane, attends a party and has too much to drink. Through the fog of her beer-goggles, she becomes loud and extroverted. While under the effect of alcohol she approaches the world as an extrovert. In the middle of loud argument about the Iraq War she thinks, in an amused but sincere way 'Those words are not my own. The beer is doing the talking!' Now we can be immediately sure that Jane remains the perspectival owner of her words and behaviour: the way they sound and the thoughts they convey (about the Iraq War) are unique to her perspective. No one else can learn of her behaviour in the same direct manner that she does. It is, rather, that she does not feel a sense of *personal ownership* towards her behaviour.

I suggest that the most natural way to analyse the absence of Jane's sense of personal ownership is to suppose that she fails to *identify with* the perspectival owner of her behaviour and hence, *as* the personal owner of that behaviour. She fails, in other words, to appropriate to her current witnessing perspective those psycho-physical cues which lend her a unique perspective on her behaviour (she may detachedly view them as mere objects of introspection). Lacking a sense of assimilation with this role, she does not approach the world through this assumed, reflexive 'filter'. As reflected in a passage by Dennett (1991, 417), it would, for example, be quite natural for Jane to feel at the time and remark later: 'I did not do that! That was not the real me talking. Yes, the words came out of my mouth, but I refuse to recognise them as my own' showing that this analysis of 'mine' in terms of 'me' fits with the way we speak and behave. Her particular thoughts and perceptions with regard to her words about the Iraq War appear to her from a unique perspective and yet she does not *identify* herself with the perspectival owner of those words. Jane feels like an owner of her extroverted experiences in only the non-identifying, perspectival sense.

Despite not feeling personal ownership towards her extroverted behaviour, Jane still feels as if the amused thoughts *about* her behaviour do belong to her. On the proposed analysis, we can say that Jane identifies with the

perspectival owner of the amused thoughts that pertain to her drunken behaviour, through appropriating to her current perspective those psycho-physical cues (including thoughts, feelings and ideas) that subtend the role 'perspectival owner of the thoughts'. Through this, she identifies *as* the personal owner of the thoughts, and so has a sense of self-identity or me-ness with respect to them. There are more extreme examples of where perspectival ownership seems to come apart from identification with a perspectival owner, such as that of depersonalisation (mentioned in Chapter 3). Such a subject, while having first-person access to the feelings in question, may report feeling as if the events 'were not happening to me'. Our analysis would suggest that while such subjects *perspectivally* own these experiences, they fail to identify with this perspectival owner as a personal 'me'.[3]

From this analysis, perspectival (or, for that matter, possessive) ownership corresponds with a notion of *perspectival* (or *possessive*) *owner* that does not imply the subject's identification of itself with this owner. To see (or hear, or feel ...) something, I need not identify with 'that which sees (or hears or feels ...)'. To think something, I need not identify with 'that which thinks', etc. Of course, most of the time we *do* identify with that which sees, hears, thinks and so forth. We identify with them in such a way that we approach the world as fully fledged *personal* owners of our experience. Whenever we harbour a sense of personal my-ness towards an X, such as a perception, action, feeling, or person, we are identifying with X's perspectival or possessive owner by appropriating the subtending basis – relevant psycho-physical cues – to our current perspective such that there is, through this apparent assimilation, the sense of *being* X's owner.

3.2.4.2. Where there is 'me', there is 'mine' (viz., a sense of personal ownership). I will now analyse and provide some evidence for the converse relation: that a subject's identification with any item (with its accompanying sense of self-identity with that item as 'me' or 'part of me') implies, for that subject, a sense of personal ownership or my-ness towards other items that are relevant to that assumed self-identity. Given our previous analysis, this sense of personal ownership towards an item will, in turn, imply the subject's sense of being a personal owner with respect to that item.

It is suggested that whenever we identify ourselves with something as 'me' or as 'part of me' – be this sports team, body image or personality – this 'broadened' subject or 'me' seems to become an integrated perspective from which *other* objects (possessively or perspectivally owned) will inevitably feel personally owned, relative to that perspective.[4] And the feelings of personal ownership, pertaining to this identity, will in turn (given our analysis so far) seem to reflect a personal owner *qua* this identity. The perspective will thus itself feel like a subject-owner, a 'platform me' from which various objects are considered to be personally 'mine'. For example, suppose that a person, 'Sid', identifies with the role of a punk rocker. In so doing, he

appropriates various thoughts, ideas and perceptions to his current witnessing perspective such that they colour the very perspective from which he perceives and thinks about the world. They may convey such things as his physical appearance (e.g., blue mohawk), the music he likes (e.g., Sex Pistols) and political ideals (anarchy, Marxism). His identity as punk rocker, moreover, seems to form part of an integrated perspective from which other items, such as his blue hair dye, feel not only possessively, but personally owned by him. His feelings of personal ownership towards the hair dye (and other punk regalia) will in turn feed into his identification *as* its personal owner *qua* punk. As long as Sid feels personal ownership towards the hair dye, 'personal owner-of-the-hair-dye' (and hence the role 'personal owner') will seem to be integrated with his overall identity as a punk rocker. Suppose that Sid renounces punkhood, enrols in a commerce degree and identifies instead as a businessman. Although remaining a possessive owner of the blue hair dye, he will no longer *identify* with this owner, and hence, no longer *as* its personal owner. He is likely to throw the dye away.

The example of Sid illustrates a case where a specific identification (with the role of 'punk') involves a reciprocal sense of personal ownership towards various items in relation to this identity (the hair dye), which in turn reflects back to an assumed identity as personal owner (of the hair dye). Indeed, it is hard to see how Sid could identify with the overall role of 'punk rocker' without assuming a reciprocal sense of personal ownership towards items relevant to this identity. The boundaries of his assumed identity-role (as punk) will of course shift with context; for example, sometimes Sid's mohawk will seem to be a part of who he is (perhaps in relation to hair dye which seems personally owned); at other times it will seem personally owned by him (perhaps when he, *qua* punk, shaves part of it off to maintain the look).

Having illustrated this analysis with reference to a specific example, it is now suggested that the analysis applies more globally; *any* identification with an item as 'me' or as 'part of me' will entail that the subject harbours a sense of personal ownership or my-ness towards certain other items that are relevant to such identification. While there is no room to provide exhaustive evidence or arguments for this claim, I will suggest that the three aforementioned modes of identification (this-ness, autonomy and consistent self-concern) all imply that the subject harbours a sense of personal ownership towards items relevant to that mode of identification. If we also accept that a sense of personal ownership implies the sense of being a personal owner (as argued in the previous section) this will show that the role 'personal owner' is the most pervasive mode of assumed self-identity and hence, central to the self we assume we are. Moreover, given that each of the three modes of assumed self-identity powerfully evidences our attribution of boundedness to the self, then so too will the sense of personal ownership that is implied by them.

A subject's identification as *this very subject*, with its dispositional feelings of value, will always involve feelings of personal my-ness towards the relevant situation, actual or imagined. If it matters (to the author) that the novel be written by him and him alone, then there will be feelings of personal my-ness towards the novel as he identifies with its author and possessive owner. If it matters that this life be lived by me and me alone, then it will feel like *my* personal life that is lived, *my* life that is in danger of replication. And if it feels like my life, then I identify as its sole personal owner. Nor is it hard to demonstrate a connection between consistent self-concern and feelings of my-ness: consistent self-concern requires that one identifies with the person on behalf of whom one is concerned. Part and parcel of this general identity is that one feels emotions of concern, emotions that are felt as personally mine, reflecting back to an apparent me who is the personal owner and bearer of those emotions.

With respect to autonomy, the story is similar. So long as I identify as the free thinker of thoughts or agent of actions, they are always felt to be personally *my* thoughts or actions. And if they are felt to be personally my thoughts or actions, then I also identify as their personal owner. The debate on free-will sometimes focuses solely on the 'could have acted otherwise' aspect, which is fine, so long as it is understood that a feeling of personal ownership is intrinsic to the feeling of free-will. Unless a subject feels that it personally *owns* its free thought or action, then the thought or action, even if genuinely free, does not feel free *to it*. In the case of tipsy Jane, for instance, her actions *qua* the extrovert feel under the control of beer rather than her own free spirit. She does not identify with the perspectival owner of her extroverted actions; she does not feel them to be hers. Even if her actions *are* free, the absence of personal ownership-feelings *towards* these actions robs her of feeling that she is their separate, autonomous agent. So with respect to these actions, she does not feel like an island of special causal power amidst a sea of blind, mechanistic causes. Rather, she feels flooded by this sea, a sea of beer whose effects wash over her character. The connection between autonomy and identification has been noted by Velleman. He claims that whenever one acts in a way that can be described as autonomous (with that feeling of free-will), the action is always rooted in a part of the personality which presents a reflexive aspect to that person, a part identified with as 'me' (Velleman, 2002, 115). In short, no behaviour with felt autonomy can issue from a part of the person that is *not* identified with as a personal owner. Insofar as the feeling of free-will makes a unique contribution to the sense of being a bounded self, it can do so only if 'the agent' also identifies as the personal *owner* of those actions.

While feelings of agency (*qua* the feeling of being a free-willed entity) clearly require a sense of personal ownership, it is not so clear that a sense of personal ownership must require feelings of agency. It should be noted that feelings of agency can arise in connection with the sheer thought of being

able to think or act freely upon a situation – one does not have to physically move about. In this way, autonomy extends to the popular characterisation of self as 'thinker of thoughts'. It is quite conceivable that one may be in a situation where feelings of agency are not activated, while feelings of personal ownership are. Suppose that Ben is lying down, enjoying a massage. It sounds odd to say that Ben identifies as an 'agent of the sensations'. It is partly because he is in a *receptive* mode that the sensations are enjoyable. But he definitely identifies with the perspectival *owner* of these sensations; he feels them to be his and he wants them to continue. It could still turn out that unless one had a *disposition* to identify as an agent, one could not identify as a personal owner. Perhaps agency is required for a sense of personal ownership in some less obvious manner. Dispositions aside, however, we can gather that the role of 'personal owner' is more immediately basic to the self than the role of 'autonomous agent'. It is the sense of personal ownership rather than agency that is easily co-definable with identification and hence, most pertinently linked to the sense of boundedness. It will hence be personal ownership rather than agency that is focused on in the later chapters that concern the ontological status of boundedness.

3.3. Integrating modes of identification into an overall conception of personality

So far, I have talked about modes of assumed self-identity (this-ness, agency, consistent self-concern and, most broadly, personal ownership) along with specific mental phenomena that, in the capacity of such modes, are appropriated to the current subject's perspective as 'me' or as 'part of me'. I will now say something about how these modes and ideas might seem integrated into a general self-conception, a conception that is reactivated and revised (even if just subtly) upon each new interaction between the subject and the world of objects. On this, John Perry offers an insight:

> Ordinarily all one's knowledge about oneself is integrated around a special sort of idea of notion of oneself that we express with 'I'. While my perception that the beer is in front of me may not require a representation of myself, the information I acquire is immediately integrated into self-attached knowledge, that I might express with 'I see a beer' or 'there is a beer in front of me'. And when I read a piece of e-mail, that says that John Perry's paper is overdue, I integrate this information into self-attached knowledge, 'My paper is overdue,' and I realize that it is me that has to get to work. I would think, 'There is a beer in front of *me*, but *I* have a paper to do.
>
> (1997, 9, online)

To express this in the vocabulary of this chapter, 'self-attached knowledge' involves both a witnessing subject (which Perry obliquely refers to as an

'agent') and an integrated idea of one's personality. The most integrated idea of all is the general idea of one's personality and history as viewed 'from the inside' or perspectivally. Given that the most generally identified-*with* role seems to be that of the perspectival owner, it makes sense to suppose that the integrated conception of one's personality is recalled in the format of either 'perspectival owner *of* personality' or 'perspectival owner *qua* personality' of some item, X. So long as we do not forget whom we identify as, this general idea (which will vary according to occasion) will be activated *every time* we perceive and interact with the world. Through identifying *with* this perspectival owner of or *qua* the personality, we identify *as* a personal owner of or *qua* this personality. Damasio (1999, 217–222) refers to this integrated personality as the 'autobiographical self' and he maintains that it is updated upon each new perception. So when Sid identifies with the persona of the punk, this persona is not an isolated thing, felt to be had by no one in particular. It becomes embedded into the content of his general self-conception. The possessive owner-of-the-hair-dye with whom he (as a witnessing subject) identifies, is not merely the possessive owner-of the-hair-dye, nor is it merely the punkish possessive owner of the hair dye. As a witnessing subject constantly encountering new thoughts and perceptions (etc.), Sid identifies with a perspectival owner (of each new thought and perception) *in the capacity of* an integrated personality; and to this continually updated personality-carrying role *perspectival owner*, the role 'punkish possessive owner of the hair dye' is appropriated. It is hence in this capacity – of the subsuming perspectival ownership role – that Sid identifies *as* the personal owner of the hair dye. And part of his subsequent embracing of the 'grey suit' – a reaction to the 'punk' – is again, not a persona to be had in isolation.

Our perceptions, however paltry, are thus not felt as isolated occurrences had by no one in particular (when they are, it is often the sign of pathology). Upon each fresh percept, the witness-subject appropriates to its current perspective not merely those current psycho-physical cues that underpin the role 'perspectival owner of this percept' but an integrated idea of 'perspectival owner *qua* personality who is the perspectival owner of this percept'. This appropriated idea becomes a filter through which each fresh percept becomes witnessed as personally mine: the idea reactivated with each new set of perspectival cues. In the words of Perry (1997, 9, online), the integrated notion of one's personality becomes *attached* to the perspective of that which witnesses each new percept, such that it does not seem merely like 'beer in front' but 'beer in front of *me*'. When this most general idea of one's perspectivally accessed personality is appropriated to one's current witnessing perspective, this greatly expands the range of potential objects that can be viewed as personally 'me' or 'mine'. Many objects that have *felt* personally owned in the past, objects as mundane as breathing, are automatically perceived as either personally me or personally mine.

Now this is not to stipulate that an overall conception of one's personality must depend upon access to a rich, personal history through first-person memory. The sense of self does not imply a sense of one's personal identity over time – at least if that timeframe is longer than 45 seconds. This is perhaps most strikingly illustrated in the case of 'David', a patient of Damasio's. As a result of a brain trauma, David lost all memory of his rich, autobiographical past and could not recall specific occurrences beyond a 45-second window (Damasio, 1999, 118). He nevertheless identified himself as a personal owner of his actions, for example, upon being pleased when winning a game of checkers. While the extent of his identification with things was of course limited, each new perception was still viewed as *personally his*, and hence, as integrated with his overall, albeit limited, conception of personality (Damasio, 1999, 113–121). Nevertheless, it should be obvious that memories of a rich personal past will, in most cases, greatly enhance one's conception of personality and hence, one's overall sense of being a unique bounded entity over time, as well as at any one time.

3.4. Identification, ownership, boundedness and *taṇhā*

I have argued that, in accordance with Buddhism, the sense of self-identity (or me-ness) and the sense of personal ownership (my-ness) are to be specified in terms of each other, each involving the process of identification. I have also argued that identification, in its various modes, offers the best evidence for a sense of boundedness. We have seen the most broadly assumed mode of self-identity to be that of the 'personal owner'; hence it provides the most pervasive evidence for the fact that we as subjects identify as a bounded, ontologically unique entity. Now I have, in the course of explaining the concept of personal ownership, been careful not to speak in a way that presupposes the existence of such an owner. I have thus spoken in the neutral terms of a subject identifying *as* the personal owner of an item (hence *assuming* itself to be a personal owner) rather than in terms of its identifying *with* the personal owner of that item, namely, *with* a ready-bound subject that is (in some way) assimilated with its role of owning the item. If we as subjects did identify *with* a ready-bound subject, then that would, in our terminology, imply the reality of such a self, viz., a personal owner. (Note that our speaking in terms of a *subject* is unavoidable when elucidating the notion of identification; should subjects turn out not to exist, then identification would also lack objective reality.)

On the Buddhist position, we as subjects do not actually identify with, and hence reflexively represent (such as through memory) a personal owner; we only *seem* to do so. A subject's identifying *with* a perspectival or possessive owner (along with associated ideas of personality, etc.) serves to mentally construct that subject's sense of being assimilated with such an owner. Hence the subject's repeated identification with a perspectival (or possessive) owner of this or that item turns out not only to *evidence* (as it seems to)

but also to *constitute* that subject's sense of being a bounded personal owner. Recall from Chapter 3 that the goal of Buddhist practice can be articulated as that of a subject losing its sense of being a personal owner. It is the *boundedness* of the self – as evidenced by a sense of personal me-ness and my-ness – that Buddhism considers primary in contributing to the self's lack of objective reality. In Chapter 8, independent evidence will be considered that suggests the Buddhist analysis to be correct: that the sense of boundedness is not in fact accounted for by an actual bounded self, but by those frequent acts of appropriation (hence identification) that *appear* to merely evidence, rather than robustly constitute, one's sense of being a personal owner.

There may be a further psychological factor that accompanies the reciprocal relation of identification and personal ownership-sense, a factor that can perhaps serve as independent evidence that such identification is taking place. Locke has notably tied the self's ownership of actions to a concomitant sense of *concern* for one's own happiness – and it is precisely this aspect which, according to Buddhism, co-arises with identification and a sense of personal ownership. In Chapter 3, we identified this aspect to be *taṇhā* or emotional investment in things being the way we would like them. To harbour a state of *taṇhā*, as we noted in Chapter 1, is to be of such a disposition that one's happiness or suffering rides upon whether states of affairs are perceived to be in line with one's desires and preferences (if not in line, then there is suffering or *dukkhā*). From our discussion in earlier chapters, we saw that Buddhism maintains that once *taṇhā* is eliminated, then so is a sense of personal ownership and self-identity, and hence, the sense of boundedness that is distinctive to the sense of self.

I finish this section by considering some preliminary evidence for the Buddhist claim that a sense of personal ownership and self-identity (hence boundedness) is associated with *taṇhā*, viz., emotional investment. If we first of all reflect upon cases where the reciprocal senses of self-identity and personal ownership (and hence a sense of boundedness) are most evidently felt, with the strongest senses of 'me' and 'mine', we will invariably find an ascription of *personal value* to the things that are identified with or felt to be personally owned. Integral to this sense of personal value will be the potential to feel happy or miserable in proportion to the degree of identification and personal ownership-sense that is borne towards that item. *This-ness*, for instance, is a major mode of assumed self-identity. Part and parcel of thisness is the value that we ascribe to our feeling of being *this very subject of experience*. Unless there was emotional investment in being *this very thing*, the thought of a replacement replica would not be upsetting (or exhilarating, if one hated one's existence). Agency is another area where mattering matters. It seems that the more powerfully we identify ourselves as the initiator (and hence personal owner) of certain actions, the more we hope that those actions will turn out as we wish. And the more hope we invest in things turning out well, the more we suffer if they do not. Hence *taṇhā*

seems to accompany a definitive sense of agency. Consistent self-concern is also value-laden; it is indeed *defined* in terms of the differential value we ascribe to 'ourselves' and to situations – feelings that are integral to *taṇhā*.

It would seem, therefore, that the more emotion that we invest in 'ourselves' or in things are perspectivally or possessively mine (whether we like them or not), the stronger our sense of (bounded) self-identity ('me') and personal ownership ('my-ness') towards those things (and vice versa). This does not mean that one must at all times be explicitly aware of one's emotional investment. As I mentioned in Chapter 1, our ascription of value to things often manifests as a *disposition* to feel various emotions in relevant circumstances, and we may be quite unaware of harbouring such a disposition until the item of *taṇhā* is in peril. We probably hardly ever think about our eyesight or breathing. It is not until something goes wrong that we suddenly realise the deep extent to which we identify ourselves with that item or its perspectival owner. People pay thousands to restore bodily functionings whose smooth running was, before the trouble, taken for granted. And if the operation turns out badly, then (provided that one is still alive and thinking!) there will be mental suffering, to the extent that one *identifies* oneself with one's body (and owner of bank account).

Emotional investment or *taṇhā* thus seems implicated in all the major modes of assumed self-identity: this-ness, autonomy, consistent self-concern and personal ownership. Perhaps, as Buddhism would claim, the feeling of differential value towards oneself and various objects is *borne out* through identification and personal ownership-sense and entirely *constitutes* our sense of a self-other boundary. As an assumed personal owner, we certainly feel that we are the source of the value-ascriptions, the home of the mattering. As James (1890, 297–298) puts it 'It [the self] is what welcomes or rejects … It is the home of interest, – not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak'. The reciprocal feelings of personal ownership and self-identity seem to tap into that deep vein of self-interest that most of us have, motivating our actions, and providing strong grounds for supposing that an *urge to be happy* is a characteristic of the self that we assume we are.

If an empirical link between identification, ownership-sense and *taṇhā* could be established, such that one could objectively *test* whether (and towards what items) one harbours a sense of personal ownership (and identification), then, as mentioned in Chapter 3, this would have positive implications for research into the nature and correlates of the personal ownership-sense (and hence, the sense of self). In Chapter 8, I will argue that Buddhism is correct in supposing that *taṇhā* and the co-arising reciprocal senses of self-identity and personal ownership not only *evidence* but also help *constitute* the sense of self-other boundedness. I will suggest that this occurs in such a manner that no *actual* bounded self need be implicated in explaining the ubiquitous *sense* of a bounded self.

4. Attribute: Elusiveness

Roderick Chisholm writes:

> The two great traditions of contemporary Western philosophy – 'phenomenology' and 'logical analysis' – seem to meet, unfortunately, at the extremes. The point of contact is the thesis according to which one is never aware of a *subject* of experience.
>
> (1969, 94)

Chisholm is referring to the feature of 'elusiveness', which, like 'boundedness', 'personal owner' and 'the witnessing subject', is central to the notion of self being explicated. It is alluded to in the selected quotations by Ryle and Damasio. 'Elusiveness' pertains to the phenomenon whereby the thing we take to be the self cannot, in its capacity as a *subject*, be directly observed by the subject itself – either through introspection or observation via the five senses. It appears that one always catches an *object* of observation such as a thought or perception, but never the original *subject* of observation, in its distinct capacity of being a subject. The subject can never, in other words, be the focus *of* its own attention; the subject is always that which *focuses* the attention. For all this, as Edey (1997) has implied, one has a definite sense of *being* a subject that is distinct from its objects, which is partly why elusiveness is puzzling. The feature of elusiveness stands behind most of the puzzles on self-knowledge, namely: given that we cannot directly observe the self as a subject, how can we know anything about it? Chisholm singles out Hume as one of the main thinkers who made 'elusiveness' famous, helping to spark the above-mentioned traditions of debate. Hume wrote:

> For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time without a perception, and never can observe anything but the perception.'
>
> (1739, 162)

It should be noted that Hume did not explicitly point to elusiveness as a feature integral to the assumed self – his exercise was not one of explicating the assumed idea of self, as we are trying to do now. Rather, his analysis *presupposed* this assumed idea of self, which featured qualities of 'uninterruptedness' and 'invariability' (still to be discussed). Upon failing to introspectively locate anything answering to these features, he sought to explain, in a manner consistent with his empiricism, the origin of such 'mistaken ideas' that we have about the self. We can thus infer that for Hume, elusiveness was not so much a feature he grappled with in and of itself, as it was

a feature that beset the very nature of his enquiry. It created for him the puzzle of where one would obtain the 'ideas' of uninterruptedness and invariability, *given* the elusiveness of these qualities (or rather, a *thing* with these qualities) to introspection. If the origin of all ideas is sensory, and all things sensory are observed as interrupted and variable, then from whence arise the contrary ideas? Any answers to this question would automatically bear upon the elusiveness issue, since the latter is woven into the problem. Common interpretation has it that Hume was ultimately dissatisfied with his empirically based solution.

'Elusiveness' and 'the possibility of self-knowledge' remain major puzzles to both analytic and continental philosophy. Many explanations have been offered – some of which deny that the self is problematically elusive[5] – but none of which are, to my knowledge, universally agreed upon. 'Elusiveness' will be returned to later in the book, when I attempt to account for this phenomenon in terms of a concept of consciousness derived from that of witnessing. At this stage of the enquiry, it is sufficient to point out that the attribute of elusiveness, evidenced by the existence of these puzzles, is central to the commonly assumed self.

5. Attribute: Unity (singularity)

> … although the things which I perceive or imagine are perhaps nothing at all apart from me [and in themselves], I am nevertheless assured that those modes of consciousness which I call perceptions and imaginations, in as far only as they are modes of consciousness, exist in me.
>
> (Descartes, 1641, 42)

I take 'singularity' or 'unity' to refer, first, to that aspect by which one unquestioningly regards all the different roles (observer, owner, agent, thinker) to be occupied by the very same *thing*, namely, the 'self' I take to be *I* or *me*. For instance, we do not reflexively assume the personal owner to be a different self from the agent, thinker or observer. All are felt to *be* the very same subject of experience – the very same *me* – whether the different roles seem to manifest at one particular time (synchronically) or over time (diachronically). This becomes apparent through reflecting on the fact that the roles are not usually identified as in isolation, but are often fluid and overlapping, with different, changing weightings. For despite this change, there remains the sense of a single underlying self that is not altered by the changing roles. Emotions are useful indicators of this point. Each emotion may involve the activation of several 'roles' at once, and there may, be more than one emotion at a time. The embarrassed person may, for instance, identify simultaneously as the personal owner and past initiator of the shameful actions, as well as the thinker of the current thoughts. At any moment, while seeming to occupy a variety of such roles, feeling a range of

shifting emotions and perceptions, one has the background impression that in spite of this affective and perceptual variation there is only one *underlying* self that is the subject of them all. Put another way, there appears to be only one point of view or perspective. The feeling is not of a tangle of different underlying selves and points of view, waxing and waning, or even of compartments of an underlying self, each dedicated to its own particular role. In fact it is very hard, perhaps even impossible, to imagine what this *would* feel like. The closest real-life examples are of those suffering with 'multiple personality syndrome' but even then, there is an arguable unity of roles within whatever 'persona' happens to be 'in control' at the time.

Another way to approach 'unity' is by looking to contemporary philosophical puzzles that relate to what is known as the 'unity of consciousness', or what Hume (1739, 163) called 'simplicity'. This debate appears to focus more on objects of consciousness than on the different assumed roles of a self, but the underlying principle is the same. A popular formulation of the unity phenomenon, relevant to this context, is as follows. One may experience multiple sensations within different sense modalities (such as the taste and texture of popcorn along with the sound and sight of a movie) and yet one automatically feels, from the first-person perspective, that they belong to the *very same* self, (in its capacity of being a subject) rather than to different subject-selves.[6] Alternatively, one could say that, at a given time, whenever a variation of perspectivally owned objects are assumed to be personally *mine*, they are also assumed to belong to a single, unified *me*. The self, in this manner, would seem to be a unified rather than a differentiated entity.

Unity must not be confused with boundedness. It is tempting to suppose that because 'unity' seems to unite or bind disparate objects to a subject's single point of view, affirmation of singularity must simultaneously affirm that subject's nature as a bounded entity, separate from other things. Unity and boundedness are, however, polar opposites: in a slogan, *unity unites, boundedness divides*. Boundedness is defined with reference to the self-other boundary; it is evidenced by identification with some, but not other items as 'me'. Unity, on the other hand, brings together objects that seem separate and divided, subsuming them under one conscious canopy. Now it may well turn out that unity and boundedness are importantly related: for example, that without unity there cannot be a sense of boundedness or identification. Boundedness thus seems, for instance, to involve a separation between a *unified* self and the rest of the world. However, this does not mean that unity and boundedness amount to the same thing. On the face of it, at least, unity and boundedness are two different aspects to the self that we assume we are.

The philosophical question behind the puzzle of unity concerns *what it is* that is responsible for this apparent unity, whether at any one time, or over time. For example, could 'synchronic unity' reduce to the mere fact *that* the objects of consciousness occur to a subject at the very same time, or is there more to such unity than synchronicity? Could the unity be, in fact, an

illusion? We will be returning to the 'unity of consciousness' later in the project. At this stage, we can note that, as with 'elusiveness', reflection on experience together with the existence of an associated philosophical puzzle, is evidence that the phenomenon of unity is a feature of the commonly assumed self.

6. Attributes: Unbrokenness and invariability

The features of 'unbrokenness' (immediate or long-term) and 'invariability' are often associated with puzzles of identity. In the debate on the unity of consciousness, identity is implicated in the phenomenon and puzzle of *diachronic* unity. The debate alludes to a commonplace impression or intuition of the *very same, unchanging consciousness or self* that unifies the various, multifarious conscious states, not only at any one time (synchronically), but also over time (diachronically), whether from one conscious moment to the next, or for the span of a lifetime. The puzzle is where we get these intuitions of unbroken sameness given that, as Hume (1739, 162) has noted, we can only detect a flux of rapidly changing perceptions when we introspectively 'enter most intimately into what I call *myself*'. This will of course tie in with the puzzle of elusiveness, since unbrokenness and invariability, should we intuit their reality, would seem to elude the usual methods of attentive observation, such as introspection and perception. After trying to account for our intuitions of unity with regards to both synchronic and successive perceptions (involving the puzzle of identity), Hume concluded in a postscript:

> But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought and consciousness. I cannot discover any theory, which gives me satisfaction on this head …
>
> (1740, 175)

In this section, it is asked whether we really do have intuitions of our self as unbroken and invariable entities. Was Hume right to make even this much of an assumption? I begin by looking to the aspect of unbrokenness.

6.1. Unbrokenness

Do we in any way take ourselves to have enduring or unbroken existence? I address this question on two levels: namely, whether we attribute unbrokenness to ourselves (1) over long tracts of time, and (2) on an experiential, moment-to-moment basis. Turning first to (1), we have to first counter an immediate objection. In an earlier section I noted that a sense of self does not require a sense of existing over tracts of time that are longer than say, 45 seconds. What, then, are we to make of the suggestion that longer-term identity is attributed to the self we assume we are? I believe the right thing to say is that while a sense of longer-term identity is not necessary to a sense

of self, it could still be (and usually is) sufficient for harbouring a sense of self and so a highly useful part of the definition. Long-term endurance, then, is still something that could qualify as important to the self we assume we are, so long as it is understood in its capacity as a sufficient rather than necessary part of that assumption.

With this proviso in place, we need not look too far for evidence of *apparent* longer-term numerical identity. First of all, there is the famous philosophical puzzle of personal identity, revolving around such questions as: what makes this person right now the *same* as that person of twenty-five years ago? Most formulations of the puzzle simply take it for granted that we have an intuition or sense of the self's long-term numerical identity; their job is to *explain* this intuition. Further evidence points to a close connection with the aspect of unity mentioned earlier. It was noted that we unquestioningly attribute all our different 'roles' (e.g., 'actor', 'owner', 'thinker', 'knower') to the *very same* subject of experience that is 'me', whether at any one time, or over long tracts of time. We do not take the underlying self to differ numerically with the varying roles and experiences with *this* self being the 'owner', *that* self being the 'actor', new selves being created every moment at the speed of thought, or upon awakening from deep sleep. The experience rather seems to be *of* fluctuating experiences, *through* different roles, *from* the one unified perspective: a perspective befitting a single selfy subject of experience.

A further source of evidence for assumed numerical longer-term identity appeals to the nature of the felt emotions. In most occasions when such emotions as guilt, fear, anxiety, disappointment, nervousness and joyful anticipation arise, their occurrence presupposes a feeling of numerical identity between the current 'self' who is feeling that emotion and a past or future 'self' to whom the emotion is related – a point noted by Canfield (1990, 19–20). Suppose that one feels guilty over a past action. It is part of the nature of that guilt to assume that the current self, feeling guilty, is identical to a past self who initiated the questionable act. Unless this identity was assumed on the part of the subject, it would be difficult, if not impossible, to feel or explain the guilt. Suppose that one is nervous of an impending exam. It is part of the nature of that nervousness to believe that some subject, namely oneself, is going to sit a future exam and perhaps not do as well as one would hope. Many occurrences of emotion presuppose an identity between what would seem to be one's current self and a past or future self. In the penultimate chapter I will argue that our very having of such emotions does not merely evidence, but also partially *constitutes* our sense of being a numerically identical self over time.

Let us now turn to (2) and consider some evidence for conscious moment-to-moment unbrokenness of self. This is the phenomenon whereby the self is somehow actively experienced as an unbroken phenomenal hum behind the changing flux of experience. I say 'somehow' because, as Hume discovered,

such ‘uninterruptedness’ (as he termed it) is not obvious. The uninterruptedness seems elusive in that it cannot be directly perceived or introspected as we perceive and introspect the usual objects of consciousness. I speak of moment to ‘conscious’ moment because there are episodes, such as deep dreamless sleep, where unbrokenness of self, even if present, is not obvious. For our purposes it is enough to argue that between such apparent lapses in consciousness, there are lengthy episodes where one is conscious and that for the duration of such episodes, the ‘self’, perhaps through its conscious or witnessing aspect, seems to exhibit the character of unbrokenness as opposed to a series of rapidly pulsing blips.

In favour of the ‘blippy’ view, Galen Strawson holds that while we might superficially attribute unbrokenness of consciousness to our experience, a little reflection will reveal, from the first-person perspective, that our moment-to-moment experience is not one of unbroken conscious existence. Rather, he says:

> My claim is not just that there can be radical disjunction at the level of subject matter [presumably, types of experience]. Switches of subject matter could be absolute, and still be seamless in the sense that they involved no sensed temporal gap or felt interruption of consciousness. It seems to me, however, that such experience of temporal seamlessness is relatively rare. When I am alone and thinking I find that my fundamental experience of consciousness is one of *repeated returns into consciousness from a state of complete, if momentary, unconsciousness* … the situation is best described, it seems to me, by saying that consciousness is continually *restarting*. There isn’t a basic substrate (as it were) of continuous consciousness interrupted by various lapses and doglegs.
>
> (1997, 422)

Strawson’s claim (which he aligns with Buddhism in a footnote!) seems unconvincing for the following reasons. First, I hold that it is logically impossible to *directly* experience what Strawson calls ‘repeated returns into consciousness from a state of complete, if momentary, unconsciousness’. Direct experience implies some concurrent conscious awareness either *of* or *with* that mode of experience – unless we are distorting the word ‘experience’ to allow for unconscious episodes (where there is nothing it is like to experience it). On the conventional usage, then, to directly experience unconsciousness is a contradiction in terms, since in order to directly experience X, one must be conscious. One cannot, therefore, directly experience episodes of unconsciousness. It is like trying to search for darkness with a flashlight.

Perhaps Strawson is alluding to *indirect* experience, such as that inferred through memory. One may, for instance, emerge from a general anaesthetic with the queer feeling that one’s consciousness has, as Strawson

(1997, 422) would put it, 'bang[ed] out of nothingness'. Perhaps this queerness is due to remembering how things were just before losing consciousness (such as the hands of the clock face) and then noting, seemingly a moment later, that things are alarmingly different – four hours have elapsed. Does ordinary waking experience seem, on reflection, to be like a series of mini general anaesthetics? I doubt it. It is true that one can have moments of distraction and reverie, sometimes being unable to actively recall a few moments of conscious activity (famously when driving a car). But this seems to confuse poor attentiveness and memory (during which there was experience, although forgotten) with nothing being experienced at all. So long as one is awake, it is highly unlikely that there is any moment when there is no object of consciousness, including those that are inattentively apprehended. And so long as there is some object of consciousness, whether direct or peripheral, we can be sure that there is consciousness or 'inner life' (whether we identify consciousness with 'witnessing' or with the stream of objects that are witnessed). It is very unusual for a person's mind, while awake, to be altogether free from thoughts or background emotions. Even if there was a temporary lull in thought processes, then at least one of the five senses will be bombarded with sensory stimuli – sounds, tastes, sights, smells and tactile experiences, not to mention an almost continuous flux of proprioceptive sensations. It is true that the sensations may rapidly arise and pass away. But importantly, there will hardly, if ever, be a moment when *no* sensory stimuli are present, or at least a moment when consciousness ceases for long enough for us to infer, as with a general anaesthetic or deep sleep, 'nothing was happening then'.

Perhaps Strawson means to assert that each object stimulates its own unique pulse of consciousness, so that there seems, on reflection, to be many numerically different but overlapping pulses of consciousness banging out of nothingness. While this may turn out to be true in theory, it still jars with ordinary, everyday experience. For it seems, to me at least, as if the objects are unequivocally presented to one and the very same consciousness – to *me*, or to *my* consciousness. That, after all, is what spawns the puzzle of unity: the puzzle of how disparate objects seem to belong to the *numerically same* consciousness, whether at any one time, or over time. If this did not *seem* to be the case, then there would not *be* such a puzzle. It would also seem to suggest that consciousness is more fittingly identified with the witnessing aspect of 'self', than with the disparate witnessed objects that come and go. I thus conclude that Strawson is quite mistaken to suppose that consciousness, from a first-person perspective, seems gappy in the way he suggests. For the reasons mentioned above, it seems to be unbroken rather than broken. It would seem, in particular, to be the witnessing aspect of self that persists seamlessly on a moment-to-moment basis.

Nor is it a merely *impersonal* witnessing that seems to persist. It has already been surmised that a sense of personal ownership towards objects of

witnessing – and hence an identity as their personal owner – is the usual case with regards to a person's conscious life. To lose the sense of personal ownership or my-ness towards such objects as ordinary perceptions, no matter how paltry, is often symptomatic of pathology (e.g., depersonalisation or anosognosia). It seems reasonable to suppose, therefore, that it is a witnessing subject *qua* personal owner – that bounded entity to whom the perceptions and other objects of consciousness seem to belong – who is assumed to persist, whether with moment-to-moment seamlessness, or over a longer span of time.

6.2. Invariability

While unbrokenness bespeaks of numerical identity, *invariability* is that aspect of the self that denotes a qualitative sameness over time (whether longer-term or moment-to-moment), a sameness that belies changes to personality and perceptions. Immediate evidence for our assumption of the self as invariable may be pointed out in the way we think and speak. When we say such things as 'I have changed' it seems that we allude to some underlying 'I' who is the unchanging bearer of the change. Nor does this seem like a mere convention; in saying this, it does *not* seem as if the essential 'me' that at an earlier time is literally different in its essence to the current 'me' at the current time. It feels, rather, as if 'I' points to an essential, unique but subtle me-making quality that belies the change and bestows numerical identity, such that 'there could only be one of me'. As Damasio puts it:

> In all the kinds of self we can consider one notion always commands centre stage: the notion of a bounded, single individual that changes ever so gently across time, but, somehow, seems to stay the same …
>
> (1999, 134)

Given that the objects of consciousness, including those identified with as 'me' qualitatively change over time, it seems appropriate to look to the aspect of witnessing when trying to pinpoint what seems invariable about the self. Because it will likely be elusive, and perhaps always occurring with attention-absorbing, altering objects of consciousness, the qualitatively neutral witnessing aspect may be understandably difficult to pinpoint, if it is indeed a component of experience. Given this, could there be further reason to ascribe 'invariability' to the commonly assumed self?

A further reason could lie in the fact that the subject-like, witnessing aspect of self, that appears to stand apart from its witnessed objects, seems also to carry its own intrinsic phenomenal feel as a self-subject (such that there seems, even if subtly, something it is like to be a self-subject). Insofar as the phenomenal feel of such a self seems subject-like rather than object-like (in seeming to elude its own attentive purview), it would not seem, from the subject's perspective, as if any phenomena being identified with are

prone to alteration. For as we discussed in Chapter 3, it is the curious nature of identification to create an effect of change-blindness towards the comings or goings of identified-with objects from the subject's conscious field. Should a change in quality be attentively noticed, then it would no longer *seem* as if the (formerly) identified-with item is a part of the subject-self at all – it would instead present as the *object* it actually is. So if the phenomenal feel of this 'me' seems to stand apart from any changeable object of awareness, then there would not seem, from the subject's own perspective, to be any qualitative parameters along which its own phenomenal character *could* be observed to vary. This could well lend the 'self' an (possibly mistaken!) impression of invariability that, in the eighteenth century, had befuddled Hume.

7. Attribute: Unconstructedness

'Unconstructedness' describes that aspect by which the self stands apart from, and is not constructed by, the kind of objects that can be perspectivally owned – thoughts, ideas, emotions, perceptions, experiences and so forth. This aspect flows naturally from everything that has been said about the self in this chapter: the fact that the self appears as a bounded, unified, unbroken, invariable, elusive *subject* of experience. In ordinary conscious states, the elusive subject always appears separate to any such objects – as that which observes, perceives, witnesses, thinks – rather than as the mental objects which are associated with such activities. In appearing to stand apart from its objects in this way, the self, with its subject-like character, purports to be of an essential nature that is not ontologically contributed to by such objects at any given time. Insofar as first-person experience leads us to forge a subject/object distinction – with the subject qualified by its selfy roles and attributes – it would seem as if there is something it is like to be this elusive, bounded self. The intrinsic phenomenal, subject-like character of this self would seem to be unborrowed from any observable objects of experience. To be sure, the self may appear to originate or 'house' such objects, as thoughts, but always in a context where the objects appear as non-identical to the self.

This remains the case, even while the subject is identifying with various objects as 'me' through appropriating ideas of them (etc.) to its perspective. Contrary to what one might initially think, appropriation does not overtly interfere with the impression of the self's unconstructedness. While appropriated ideas are still, ontologically speaking, *objects* of consciousness – of the sort that *can* be perspectivally owned and attended to – their status as such objects becomes, for reasons already mentioned, *effectively invisible* to the identifying subject. If they did not become invisible, but still appeared as overt objects to the subject, then they would not have been appropriated to the perspective of the subject (they could be separate thoughts *about* the appropriated idea, for instance). When 'the self' confronts the world as

'other', any appropriated idea will always be *qua* a unified elusive subject, never *qua* a separate observed object. This curious shrouding effect that appropriation has upon the selected object-ideas thus does nothing to impede (and perhaps even enhances) the subject's impression that as a bounded self, its ontology is free from any potential contribution by those appropriated object-ideas. The subject hence retains the impression that the self, the thing it assumes it is, is unconstructed.

The subject's appropriation of ideas to its perspective does not imply that the self is actually something that is constructed, in any essential way, by those appropriated ideas. While appropriation serves as the best *evidence* for the fact that we assume our selves to be bounded – evidence that does not seem to interfere with the self's unconstructedness – it does not prove that the boundedness is, in fact, *constituted* by such appropriations. From the first-person perspective, indeed, it would seem as if a ready-bound self is identifying with various items, rather than as if the identifications are constituting the very boundary between self and the world. Buddhism holds that appropriation does constitute the sense of boundedness – but this is moving away from our concern of characterising the self we assume we are, to a different concern of how that assumption is generated.

There are further matters to be clear about. By ascribing unconstructedness to the self that we purport to be, I am not saying that the self purports to be the type of thing that could, at a given time, exist in isolation from all objects of consciousness. Nothing about the assumed self seems to preclude the possibility of the self being something that at any given time will necessarily co-exist with its overtly observed objects. Perhaps consciousness with a sense of self is inherently dualistic, such that, whenever there appears to be a self, there must also appear to be objects that are observed by the self. What is being claimed, rather, pertains to that very sense of dualism: the unbridgeable gap that is felt between self (with its intrinsic phenomenology) and object in a given state of consciousness, at a given moment.

Nor am I saying that unconstructedness *per se* amounts to boundedness (although unconstructedness *of the self* implies boundedness). Unconstructedness simply pertains to the fact that the self *qua* witnessing subject does not seem constructed by the kind of mental objects that are witnessed, just as diffuse light does not seem reducible to the objects it illuminates. Boundedness (e.g., as evidenced through identification) seems to add a definitive although elusive border of 'me' around this witnessing, as a cloud of dust might seem to add shape to the diffuse light. Just as the collective light-and-dust can seem to offer a unified source of illumination for other objects, the subject-and-identified-with-item can seem, from the first-person perspective, to be a unified source of observing, perceiving or thinking about other objects. Unconstructedness pertains *strictly* to that 'illuminating' or witnessing aspect by which the subject – bounded or not – does not seem to be constituted by any observable mental objects.

Conclusion

It has been argued in this chapter that most people do indeed take themselves to be self-entities of the kind alluded to by both Buddhism and Western philosophers. That is, most people take themselves to be a witnessing or conscious subject that occupies the roles of observer/witness, owner (perspectival and personal), free-willed agent of actions and thinker of thoughts, and happiness-seeker/avoider of suffering. People also assume this same self-entity to bear the attributes of ontological uniqueness/boundedness, elusiveness, unity, (short- or long-term) unbrokenness and invariability, and unconstructedness. We already know that Buddhism denies objective reality to this self as a whole. As it happens, many Western philosophers have also denied objective reality to this self and in the next chapter we look to some such evidence. Prior to this, I will investigate what it means to deny reality to such a self and whether such a denial could amount to the same thing in Buddhism and the West, given different background assumptions.

5
How Do We Construe 'The Self Lacks Reality'?

Introduction

Having determined that there is a commonplace *assumption* of self, we can move onto the question of whether we are *indeed* such a self: does the self we assume we are actually exist? Is it real? The first step of this enquiry will be to analyse what it could *mean* for this self to lack reality: the goal of this chapter. I will suggest that the self, if it lacked reality, would lack reality by virtue of its being *constructed* and *illusory* (concepts that I spend some time explaining). The analysis will then be put to the test through its application to positions of some key Western thinkers who have denied reality to the self. By referring to Western thinkers who regard the self to be constructed and illusory, we further demonstrate the relevance of this project to Western concerns: casting the self as constructed and illusory is not peculiar to Buddhism. The proposed analysis of *construct* and *illusion* will also be flexible enough to convey major *differences* between the Buddhist and Western accounts of 'no-self'. It was noted in Chapter 3 that while Buddhism (on my analysis) regards boundedness (grounded in the role *personal owner*) as the main reason for the self's illusory status, it does not deny reality to some further ascribed features such as unity, unbrokenness and invariability. By contrast, it will be seen that the Western counterparts (as with the *orthodox* Buddhist readings) *do* deny reality to these and other features by regarding them as constructed and illusory. There will hence be significant divergence in the *extent* to which the reality of self is denied by Buddhism (interpreted) and the West: a divergence not emphasised in the comparative literature.

1. What does it mean to deny reality to the self?

If 'reality' pertains simply to the metaphysical standard by which things are granted or denied existence, then we may have a difficult time finding an agreed-upon standard by which the self could be said to lack reality.[1] We have already seen that the metaphysics of Buddhism, in relation to which the self

is supposed to lack existence, involves the unconditioned *nibbāna*. In Western philosophy, whose metaphysical standards would generally *preclude* the reality of *nibbāna*, denying existence to the self may mean something quite different. For example, when Richard Baron (2000) says that the self is unreal, the standards by which he construes the non-existence of self are those informed by what is sometimes known as the 'scientific image'. On the scientific image, the sum total of things regarded as real are things that can be revealed through careful scientific method. Such methods, says Baron (2000, sec 2–3) reveal no such entity as an observing, thinking self with this-ness and autonomy. Western philosophers who reject this elevation of the 'scientific worldview' may nevertheless deny the self reality because they think that anything that truly exists must be an object that can be the focus of attention. Being elusive to attentive purview, the self falls outside the scope of reality; such a line seemed to be taken by David Hume (1739).

Given the divergent metaphysical systems that can inform one's favoured standard of reality, can there be any convergence in the way that different thinkers, regardless of background metaphysics, deny the self reality? I believe that there can. The self has sometimes been claimed, in both Buddhism and the West, to be *illusory* and at other times as fictional or *constructed*. While these claims are not equivalent, it will nevertheless turn out, in the case of the self, that its constructed status would make it illusory. Once we articulate what it could mean for the self to be illusory and constructed, it will become clear how the self could be denied reality in a way that stays neutral on the background metaphysics.

1.1. What is an illusion?

I begin by analysing the concept of *illusion* and then, in a later section, that of a *construct*. The possible application of these concepts to the self will become apparent as discussion proceeds.

Most generally, an illusion involves a conflict between appearance and reality. Something, X, appears to be the case, but there is something about X that does not reflect reality; it *misleads* the person to whom it appears. In other words, X *purports*, through the appearance, to exist in a particular manner, when X does *not really* exist in that purported manner. More formally:

> When X purports (through a medium of appearance) to exist in manner F, to person P, X-as-F is illusory when X does not really exist in manner F.

For reasons that will become apparent, I have kept this formulation very general. It will cover not only standard perceptual illusions (such as the Müller–Lyer Illusion) but also delusions (such as that of the man who believes he is being monitored by aliens) and hallucinations, such as that of the pink elephant squatting on the Sydney Harbour Bridge (did you see it too?). I now elaborate further upon the formulation.

All illusions (including delusions and hallucinations) will involve appearances to the perspective of a conscious *person* (a psycho-physical entity to whom selves are normally, even if mistakenly, ascribed). Indeed, unless we significantly altered the ordinary meaning of the term 'illusion', it is not at all clear how an X-as-F could count as illusory and yet X not appear (or be disposed to appear) in some (deceptive) way to a conscious *person*. (It should be noted also that while I actually think the term 'subject' rather than 'person' to be more apt in this definition, I do not wish to beg questions at this stage by presuming that subjects exist. Nevertheless, it may in the end turn out that reference to the subject is unavoidable when specifying the nature of illusion).

The medium of deceptive appearance will borrow its cloak from either the five perceptual senses, or that notion of sense *qua* conscious impression discussed in Chapter 1 (e.g., 'a sense of danger'). So by the formulation 'X purports to exist (through a medium of appearance) in manner F' I mean that X appears to a person in such a way that X seems to 'wear on its sleeve', via the medium of such appearance, a manner of existing that goes beyond its mere appearing-to-P. So to person P, it will seem as if appearance-independent reality – conveyed by the term 'manner F' – *objectively grounds* the appearance of X.

In this way, appearances of X-as-F are cognitive, openly conveying a message about X's manner of existence – a message that turns out to be false (viz., 'X does not really exist in manner F'). For example, suppose that X is 'two lines of unequal length' and F is 'existing independently of P's perspective'. In the case of the Müller–Lyer Illusion, a message of X-as-F is conveyed to the person; there is the appearance of there *really being* two lines of unequal length. From the perspective of P, the appearance of unequal lines (X) purports to be grounded in a perspective-independent state of affairs (manner F). What the appearance conveys – X-as-F – will be illusory because the independent situation does not *in fact* ground X as existing in manner F (as would *seem* to be the case) but rather shows X to be existing in a manner that is not-F.

A person harbouring an illusion need not be fooled by it to the extent that she actually *believes* X to exist in the purported manner, F. Those familiar with the Müller–Lyer Illusion will not actually believe the lines to be of objectively uneven lengths. Yet the lines will still visually appear to them *as if* they were objectively uneven, such that were they not to recognise them as illusory, there would be psychological pressure to believe the lines to be uneven. It is in this capacity that the phenomenon gets classed as a standard optical illusion. Similarly, someone on hallucinogenic drugs who 'hears' a chorus of angels as purporting to originate from heaven (as opposed to from their mind) will not, should she know this to be the effect of the drugs, believe that she really *is* hearing a heavenly chorus. The heavenly chorus (the content of a hallucination) will nevertheless be illusory, purporting to

originate from heaven – and hence independently of her perception – when it does not actually exist in this manner (it exists only 'in her mind').

Sometimes the illusion will cut deeper into one's cognitive set so that the very appearance of X involves central *beliefs* about X's mode of existence. The schizophrenic who hears voices in his head and is ignorant of his condition is likely to completely buy into the objective reality of these voices, believing, for example, that the voice is of God telling him to go on a mission. When beliefs as major as this are drawn into the illusion, the phenomenon tends to be classed as a *delusion*. The voices from God nevertheless qualify as illusory on our definition, since they purport to exist in a manner (namely, as really from God and hence independent of perception) in which they do not actually exist. It is just that the 'mechanism of purporting', as it were, draws upon a deeper set of beliefs in the subject's cognitive set.

While the delusion of voices from God will involve an auditory impression, other delusions may be presented through a medium of appearance that draws not upon the five perceptual senses, but upon a sense *qua* conscious impression. Suppose that Psyaan is gripped by the conviction that aliens are watching him. He has not been presented with any perceptual-like cues that bespeak of this impression, yet he still has a distinct *sense* of being watched by aliens. While the impression is of being watched by aliens, the watching aliens need not – and do not – have the ontological status they purport to have, that of existing outside of Psyaan's mind and causing him to have such experiences. The fact that they purport to have such status, from the perspective of Psyaan, renders the watching aliens illusory.

Now in its normal usage, the term 'delusion' has connotations of involving a person's deeper beliefs – and so of being 'doxastically committed'. Delusions also imply a cognitive malfunction or abnormality, such as in the case of Psyaan. The medium of appearance in a delusion may or may not be perceptual (compare believing in 'voices from God' with believing in a sense that aliens are watching). By contrast, the term 'illusion', at least in its more restricted usage, has connotations of being perceptually oriented but doxastically uncommitted (such that it may or may not incur a deeper belief) and of *not* implying malfunction or abnormality. Indeed, if one *failed* to optically appreciate the Müller–Lyer Illusion, then *that* would be a sign of abnormality or malfunction! Hallucinations, by further contrast, are usually presented through a *perceptual* medium of appearance (e.g., auditory); they are doxastically uncommitted and involve abnormality.

When it comes to articulating how the *self* could possibly be illusory, it does not seem to fit any such camp. The self, if lacking reality, would be like a delusion insofar as the self-sense would involve doxastic commitment: the conscious impression conveying a reflexively assumed *belief* in the self's existence, a belief that turns out to be false. And yet unlike a delusion, this reflexively assumed belief will not be considered, at least in the West, to

involve a cognitive malfunction or abnormality. (Even Buddhists would agree that it is statistically very normal.) Nor would the self obviously class as a standard illusion since, like the sense of being watched by aliens, the medium of its appearance would be non-perceptual. And for this reason also, it could not be a hallucination. The broader definition of illusion that I have proposed is thus needed to properly capture the way in which the self might, plausibly, count as an illusion. We can hence surmise, on this broader definition, that should the self be illusory, it would not exist in the manner that it purports – via the sense of self – to exist. (I later specify just what this purported manner would amount to.)

Before going further, I need to avert a possible confusion by pointing to an ambiguity in the way that the term 'illusion' can plausibly be deployed. We need to distinguish illusion *qua* the appearance-of-(X-as-F)-to-P (the vehicle) from Illusion *qua* X-as-F (the content, viz., thing *conveyed* through the appearance-of-(X-as-F)-to-P). While our stipulation defines the term 'illusion' in the latter sense, the former sense is also common. With the Müller–Lyer Illusion, for instance, the very visual impression of two objectively uneven lines to a person is commonly termed 'an illusion'. But the illusion *qua* visual impression is not the thing whose existence is in question. Rather, the thing whose existence is in question is the content of that appearance, namely, the illusion *qua two objectively uneven lines*. Indeed, unless the illusion *qua* appearance-vehicle (to P) existed, there could not *be* the illusion *qua* content-of-appearance, viz., two objectively uneven lines. A similar story applies to the other kinds of illusion. The term 'illusion' can refer to the auditory impression of purportedly external voices, such that the subject can be said to be harbouring such an impression, or be in the grip of an illusion. In this manner of speaking, the illusion exists as a real something that is 'had' by the person. But we can also say, as we have predominantly been doing in this discussion, that the *voices*, since they purport to exist externally, are an illusion (or illusory). And regarding the illusion of self (should the self be an illusion), we already know that the actual sense of self, which would serve as the vehicle/appearance for the illusory content, would, relative to the self, be real enough. Indeed, unless there actually *was* a sense of self, then there could not be a conveyed *self* that may turn out not to exist in its purported manner. So unless I stipulate otherwise, I shall continue to use the term 'illusion' or 'illusory', to talk about the X-as-F (whose existence is in question) rather than the appearance of X-as-F (whose existence is not in question). We should also at this juncture note that the term 'content' being used in this context – and indeed throughout this project – is not equivalent to that of 'object'. That is because the content of an appearance, such as the self, need not be accessible, in principle, to the subject's attentive purview – although it must be accessible to its wider experiential purview (e.g., peripheral).

1.2. The self as construct and illusion

Before considering how the self's purported manner of existence may conflict with its actual ontological status, it will be useful to first analyse in more detail how such a situation may apply to phenomena that are not the self. From this discussion the concept of a *mental construct* will be introduced; then the analysis will be applied to the self.

In the examples of illusion we have looked at, there has been an implicit standard of reality ('manner F') by which (a) the item in question has, from the perspective of the person, purported to exist and (b) the item is deemed, from a perspective independent of the person, not to exist. This independent perspective has, in such cases, involved a reality that is literally in the world *outside* of the appearance-vehicle in which the (illusory) content is harboured.

With the Müller–Lyer Illusion, for instance, the standard of reality by which the content <two objectively uneven lines>[2] seems optically real to the observer, is a reality that seems, from the observer's perspective, to exist *outside of* its actual appearance to her. It purports to have a reality that will meet any measuring criteria that would prove it to be an objective, observer-independent, fact. It is by reference to this appearance-independent standard of reality that <two objectively uneven lines> proves to be an illusion, not existing in the manner that it purports, from the person's perspective, to exist. Likewise, the person who 'hears' voices from God presumes (with greater doxastic depth) that the voices-of-God exist *outside of* the mere appearance of this voice to his perspective. This purported observer-independent reality serves as the standard by which the voices-of-God get judged (by outsiders) as illusory: as not existing in the manner that they purport, from the person's perspective, to exist.

When such an appearance is not grounded in external reality, an alternative explanation for the appearance is required which does not, as part of its explanation, make appeal to the item purported to be real (via the appearance); for example, *two uneven lines* or *being watched by aliens*. The explanation will instead include such factors as neuro-chemical imbalance, but importantly, it will always appeal to cognitive and conscious components that directly contribute to the appearance of the entity in question, whether the appearance be in the medium of perceptual-like cues or conscious impressions. Such an explanation will hence appeal to such factors as memory, imagination, sensations, emotions, thoughts and perceptions, with perspectivally owned objects of consciousness (whether occurently attended to or not) figuring in the explanation. In this way, we can say that the content <two objectively uneven lines> or <being watched by aliens> is nothing other than a *mental construct*, not existing beyond the subjective impression in which it appears. The illusory status comes about from the fact that it purports to have an *unconstructed* status, that is, to exist

beyond the appearance, of which it is the content. The unconstructed status is the 'manner F' in which an item is purported to exist. Hence:

> When X purports (through a medium of appearance) to be unconstructed (existing beyond the appearance of which it is the content), to person P, X-as-unconstructed is illusory when X is really constructed.

In view of this, we can now define the notion of a mental construct or a *constructed X* (and its contrast-class of *unconstructed X*) along with the notion of *illusion* as it stands in relation to this definition (with the appearance to P made implicit):

> D1: Constructed X: X is a construct iff X does not exist beyond the appearance, of which it is the content.
>
> D1: Unconstructed X: X is unconstructed iff X does exist (and is causally grounded) beyond the appearance, of which it is the content.
>
> D1: Illusory X: X is illusory iff X is a construct and yet purports to be unconstructed. That is, X does not exist beyond the appearance, of which it is the content, and yet X purports to exist beyond the appearance, of which it is the content.

Suppose I 'hear' a voice. The voice is the content of an auditory-like appearance. I appear to hear a voice and I assume that the voice 'heard' has an existence independent of its manner of appearing to me. Suppose that the voice does in fact exist in its purported manner and is independent of me, *represented by* its auditory appearance. In our analysis, the voice turns out to be unconstructed and non-illusory, as we should expect. Suppose that I now 'hear' a voice that sounds just like the first and which also purports to exist independently of its auditory appearance. The voice, however, is a figment of my imagination, and is hence a construct that does not exist beyond the auditory appearance, of which it is the content. Its existence actually depends upon input from cognitive factors contributing to the hallucination – factors which do not include an external voice in their ontology. In our analysis, the voice will be illusory (as we should expect) since it purports to exist in a manner (namely, as an external, unconstructed voice) in which it does not actually exist (it is instead a mental construct).

I used to think that the analysis allowed for constructs that are not illusory, but now I am not so sure; it is actually quite difficult to find an example. While thoughts can no doubt contribute to mental constructs, a mere thought about an object (such as a chair) does not itself count as a construct since it does not involve an immediate appearance that is cloaked in the medium of conscious or sensory impression (such as a sense of aliens

watching, or apparition of a chair). And coloured afterimages, while passing the apparition test, do not seem to have enough content or 'aboutness' to separate them from the medium of appearance itself. If a yellow afterimage *was* to present in such a capacity as 'colouring' a white wall, then it would become an optical *illusion*. There seems to be no intermediate position where the afterimage has both content and is not illusory. Perhaps there are no proper cases of constructs that are not illusions; perhaps the terms 'construct' and 'illusion', as I have defined them, are co-extensive.

We are now in a position to see if our analysis applies to the self. If the above definitions are applicable, we should be able to envisage a possibility whereby the self, while purporting to be unconstructed – existing beyond the appearance, of which it is the content – is in fact constructed, viz., not existing beyond the appearance, of which it is the content.

But this does not, upon reflection, seem like a possibility at all. The self is a peculiar kind of entity, in that it seems to represent its own objective existence both transparently and reflexively by simply *being* the subjective entity that it is. In this way, the self does *not* purport to exist outside of its very own appearance (viz., the sense of self) – and yet unlike an afterimage it has, via the sense of self, a proper content or 'aboutness'. So while on one hand, the sense of self does convey the message <conscious, witnessing, bounded, elusive, unified subject (etc.) who is owner, actor, thinker (etc.)>, on the other hand, there does not seem, phenomenologically, to be a gap between the message conveyed (<the self>) and that which conveys the message (the sense of self). The self seems to be its own ontological ambassador, with the sense of self appearing to be none other than the self. Just as there is no felt gap between the subject *per se* and the self (as discussed in Chapter 1) there seems to be no felt gap between the self one purports to be, and the sense of being that self. What it is like to be a conscious self who thinks thoughts (etc.), seems no different from what it is like to have a *sense* of being a self who thinks thoughts (etc.). The self we assume we are simply *is* a conscious entity with certain roles and attributes, its lived, internal what-it-is-like-to-be-a-self-ness (in relation *to* its observed objects) packed into its ontology, so to speak. To have a sense of being this entity does not seem to require an additional meta-sense, such that there seems something it is like to be an entity, of which there is something it is like to be that entity. In short, the self purports not to exist outside of the very appearance (the sense of self) of which it is the content, since the sense of self seems to reflexively represent the self's existence as if it *were* that very self. So the self comes out, on the above definitions, as purporting to be constructed – not unconstructed, as we should expect.

There is something that unites those items to which the above definitions of *constructed*, *unconstructed* and *illusory do* successfully apply – something that could shed further light on why the definitions do not apply to the self and what it would actually take for the self to be unconstructed. Whether

constructs or not, <singing voices>, <being watched by aliens>, <objectively uneven lines> and <pink elephants> are all *objects* of consciousness; all are accessible to attentive purview. Perhaps, then, the above definitions fail to apply to the self precisely because the self *does not* purport to be an object, but a (unified, bounded, etc.) *subject* of consciousness, with its content bound up in its mode of presentation. The mode of presentation, for any subject-like entity, must be one that purports to stand it apart from any objects that could appear in its 'field' of witnessing, objects such as thoughts, perceptions and imaginations. In purporting to stand apart from such objects that could appear to it, there *could* be no felt gap between the self that presents as a unified, bounded subject of experience and the *sense* of being that entity. If there was a felt gap, such that the self appeared as an attentively observable object to an observing subject that sensed it – *then it would no longer qualify as a self*. Part and parcel of selfhood is that it presents itself as being *integral to* the observing subject – of which there is something it is like to be that observing subject – never as an observable, separate object. It is in this capacity (talked about in Chapters 3 and 4) that the self is presented as being an unconstructed, subjective entity – unconstructed by any object that is *capable of being* perspectivally owned by it – objects such as thoughts, emotions, feelings, imaginations and so forth.

In view of these reflections, it can be seen that the term 'manner F', as it features in the opening definition of illusion, is ambiguous. The ambiguity centres around how to interpret the idea of an X as being 'perspective-independent' or 'existing independently of appearance'. It can either mean, as it does in the case of objects, that X is grounded in a reality that is literally *outside* of the appearance or perspective to P, or it can mean, as in the case of the self, that X is grounded in a reality that veridically *underpins* the appearance or perspective to P. For the self to be underpinned by this kind of reality it must be – as it seems to be – unconstructed by any perspectivally ownable objects that could appear to it as 'other' (namely, as a focus of attention). In view of this, we can now propose the following additions to our definitions such that they can accommodate the subject-like self as well as objects:

> D2: Constructed X: X is constructed iff X is the content of an appearance, by which X is contributed to by such objects as thoughts, sensations, perceptions and emotions AND X does not exist beyond the appearance, of which it is the content.
>
> D2: Unconstructed X: X is unconstructed iff EITHER X is the content of an appearance, by which X that is NOT contributed to by such objects as thoughts, sensations, perceptions and emotions, OR X does exist (and is causally grounded) beyond the appearance, of which it is the content.
>
> D2: Illusory X: X is illusory iff X is constructed and purports to be unconstructed. That is, X is illusory iff EITHER X is the content of an appearance,

by which X is contributed to by perspectivally ownable objects and yet X purports to be the content of an appearance, by which X is not contributed to by such objects, OR X does not exist beyond the appearance, of which it is the content, and yet X purports to exist beyond the appearance, of which it is the content.

These definitions allow us to express the way in which the self, as well as objects, can be constructed, unconstructed or illusory. If the self is a construct, then it will exist as the content of an appearance, by which it is contributed to such objects as thoughts, emotions and sensations, and it will not exist beyond *this* appearance, of which it is the content. Because the self purports to be *unconstructed*, that is, to be the content of an appearance, by which it is *not* contributed to by such objects as thoughts and so forth, then the self, if actually constructed, will be illusory. The self-as-unconstructed (viz., the self) will be purporting to exist in a manner in which it does not actually exist – as something unconstructed when it is in fact a construct.

It could hence conceivably turn out that the unconstructed entity we deeply assume we are, does not have the ontological status that it purports, from the first-person perspective, to have. Instead of being underpinned by an actual unconstructed self-entity, the sense of self would be underpinned by factors that do not appeal to 'the unconstructed self' in their ontology. Such factors would include perspectivally ownable objects like thoughts, emotions and sensations – those very items to which the self seems to stand in opposition! At pain of oversimplification we can thus say, on such a scenario, that rather than the self thinking the thoughts (as it seems to do), the thoughts would be helping to think the self.

1.3. An intersection of agreement for those who deny reality to the self

Our analysis provides a clear intersection of agreement for thinkers who deny reality to the self, despite different informing metaphysics. They can all agree that the self fundamentally lacks reality because it is an illusion, purporting to exist in a manner in which it turns out not to exist. They can all agree that the self purports to exist as something that is unconstructed by thoughts, emotions and so forth (*unconstructedness*, after all, is essential to the self's definition). They can all agree that the self's purported manner of existence, as unconstructed, conflicts with its actual manner of existence, as constructed. They can all agree that the appearance-with-content (namely, the sense of self) will be accounted for not by an unconstructed self (which would seem to account for it), but by cognitive components that include such objects as thoughts, ideas, perceptions and emotions. In short, there can be enough agreement upon the manner in which the self fails to exist – through its being deemed as constructed and illusory – to make general sense of the claim that the self lacks reality.

This convergent way of denying reality to the self still manages to accommodate the divergent reasons for which the self can be said to lack reality. The divergences will be reflected in varying accounts of how and why the self is constructed (and hence illusory). For example, someone taking a Humean or Jamesian line (soon to be discussed) may claim that the self comes to be constructed through the fact that its unity and identity (which includes unbrokenness and invariability) has no reality beyond a bundle of discrete thoughts and perceptions (etc.), which contribute to the *appearance* of such unity and identity. Should such an account be correct, then this, together with the fact that the unity and identity are *purported* to be unconstructed features of the self, would render the self, as a whole, to be a construct and hence illusion. Someone taking a Buddhist line, on the other hand, may say that the self comes to be constructed through the fact that its boundedness, through the process of identification, is a product of nibbānic witness-consciousness appropriating various ideas to its perspective. They will claim that through the mechanism of appropriation, the impersonal, nibbānic witnessing comes to be felt as a personal bounded self, rendering the boundedness (because it depends upon contribution from thoughts (etc.)) a construct. Should such an account be correct, then this, together with the fact that the boundedness is *purported* to be an unconstructed feature of the self, would render the self, as a whole, a construct and hence illusion. The following schema provides a threefold way of thinking about the scenario of *self as constructed and illusory*. It will be useful to reflect upon this schema when looking at evidence from Western philosophers who deny reality to the self.

1. Through various roles (e.g., as thinker of thoughts, owner of experiences, initiator of actions), we take ourselves (*qua* self) to be a conscious subject-entity, which is ontologically separate from, hence unconstructed by, thoughts, perception and other experiences. In reality, such a self is not ontologically separate from the thoughts (etc.), but is the content of an idea that is created, at least in part, by our thoughts, perceptions and so forth. Thus the self does not precede or create the thoughts (etc.); rather our thoughts (etc.) go towards creating the idea of it. In this way, the fundamental duality between a thinking, perceiving, bounded subject on one side and its apprehended objects, thoughts and ideas on the other, is constructed and hence illusory. This has bearing on the various roles and aspects as they are applied to self.
2. We take our thoughts (etc.) to be owned – perspectivally and personally – by a self, when in reality they are not owned by such a self. The idea that we, the self, own our thoughts and perceptions (etc.) is caused at least in part by the edifice of thought and perception (etc.) that comprises the sense of self, rather than by a thought-independent owner, the self. Similar considerations apply for other roles such as actor and thinker. We take the self to be the initiator of actions and thinker of thoughts.

In reality, there is no self playing these roles; 'actor' and 'thinker' are, at least in part, a fictitious invention *of* the thoughts (etc.). Neither is there a *bounded* witness-self who knows the experiences (although Buddhism construes witnessing to be real enough).

3. We take ourselves to be a bounded, unbroken and invariable, unconstructed entity, which, through various roles, *unifies* our thoughts and experiences as belonging to and issuing from *it*. In reality, there is no such self, but only a flux of thought and perception along with mental faculties such as memory and imagination. The Buddhist account also includes witnessing, which is construed as unbroken and invariable, a source of the apparent unity. But importantly, there is no room in this picture, whether painted by East or West, for an entity described as 'the self' that serves to unify the thoughts. If there is a genuine principle of unity, then this principle is not grounded in the self-entity.

2. Western thinkers who deny reality to the self

It would seem that Buddhism is not alone in denying the self existence. I will now test out the analysis by considering some leading Western thinkers who regard the self not to exist. From their quotations to follow, it will be evident that while agreeing with Buddhism that the *self as a whole* is constructed (and hence illusory), these Western thinkers provide a very different story as to how the self is constructed. The features of unity, unbrokenness and invariability are usually singled out as the *primary* reason for why the self, as a whole, is to be regarded not to exist. These features, in and of themselves, are considered to be entirely constructed (and hence illusory).

We begin with Hume (1739). Let us first recall the evidence that Hume was indeed concerned with the notion of self we have been discussing all along. When he writes: 'But the self or person is not any one impression, but that to which our several impressions and ideas are supposed to have a reference' (Hume, 1739, 161–162), it seems he is capturing the relevant notion of self, viz., an entity that *has* thoughts, ideas and perceptions, but is not identical to them. Hume (1739, 163) also ascribes 'simplicity' to the self, a principle of unity by which a subject's diverse thoughts and perceptions are perspectivally owned by a single entity. Most famously, Hume's (1739, 163) discussion pre-supposes a common ascription of *identity* to the self, by which we are compelled to 'suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives'. Now the actual state of affairs, according to Hume, is quite contrary to how things appear. Instead of there being an unbroken, underlying entity which unites the varying perceptions and accounts for their identity, there is merely:

> a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement

> (Hume, 1739, 162) …. There is properly no simplicity in it [the mind] at any one time, nor *identity* in different, whatever natural propension we may have to imagine that simplicity and identity (Hume, 1739, 163).

He goes on to argue that:

> The identity which we ascribe to the mind of man is only a fictitious [viz., constructed] one, and of a like kind with that which we ascribe to vegetables and animal bodies. It cannot therefore have a different origin, but must proceed from a like operation in the imagination upon like objects (Hume, 1739, 168). … identity [and simplicity] is nothing really belonging to these different perceptions, and uniting them together, but is merely a quality which we attribute to them, because of the union of their ideas in the imagination when we reflect upon them (Hume, 1739, 169).

Hume clearly regards the appearance of the self's unity and identity to be underpinned not by factors that include *actual unconstructed* unity and identity – whether the identity be construed as immediate or long-term – but by mental factors such as 'the union of their ideas in the imagination when we reflect upon them'. The self, on Hume's theory, is a mental construct *in virtue of* its constructed unity and identity. The illusion is hence that of an unconstructed entity: a self, whose features objectively underpin our sense of unity and identity. It is an illusion because no such principle underpins this impression; there is only a diversity of rapidly fleeting perceptions acted upon by the memory and imagination (not unlike the orthodox reading of Buddhism on this matter!)

Moving on now to William James, recall that in *The Principles of Psychology*, James writes:

> … [the self] is the source of effort and attention, and the place from which appear to emanate the fiats of the will … Being more incessantly there than any other single element of the mental life, the other elements end by seeming to accrete round it and to belong to it. It becomes opposed to them as the permanent is opposed to the changing and inconstant.
>
> (1890, 297–298)

This 'spiritual self' says James (1890, 301), is empirically known to us through bodily feelings – especially around the head and throat – although we may not be explicitly aware of this fact. But he concedes that such feelings are not enough to account for what it is that unites (I would say, 'seems to unite'), both synchronically and diachronically, the diversity of experience we call our 'own'. In the following passage, James identifies the roles of (perspectival and personal) *owner*, *knower* (witness), *thinker*, *initiator of actions* and the

attributes of *unity* and *identity* with a section of the 'stream of consciousness', namely, with whatever *thought* one may have at the present moment:

> … the real, present onlooking, remembering, 'judging thought' or identifying 'section' of the stream. This is what collects, – 'owns' some of the past facts which it surveys, and disowns the rest, – and so makes a unity that is actualized and anchored and does not merely float in the blue air of possibility (James, 1890, 338) … It is the Thought to whom the various 'constituents' are known. That Thought is a vehicle of choice as well as of cognition; and among the choices it makes are these appropriations, or repudiations, of its 'own.' (James, 1890, 340) … The *I* which knows [knows the empirical aggregates] cannot itself be an [p. 401] aggregate, neither for psychological purposes need it be considered to be an unchanging metaphysical entity like the Soul, or a principle like the pure Ego, viewed as 'out of time.' It is a *Thought*, at each moment different from that of the last moment, but *appropriative* of the latter, together with all that the latter called its own (James, 1890, 400–401).

Like Hume, James (1890, 339)[3] attempts to fully account for 'that appearance of never-lapsing ownership for which common-sense contends' without positing an unbroken, invariable metaphysical 'Owner' that would do the work of the self (James, 1890, 338–339). His solution (as with Hume) is to propose that such a self amounts only to content that is presented by mental objects (the 'judging thought'), as opposed to something substantial that underlies and precedes the thoughts.[4] When he rejects the idea that the self offers a grounding for the impression of unity, unbrokenness and invariability (something Buddhism would agree with) James also, like Hume, rejects the idea that those impressions could have *any* grounding in unconstructed reality (setting his account apart from that of Buddhism). He attributes their impression entirely to the 'judging thought' whose nature it is to change from one moment to the next. Through this move, James' account renders these and other features of the self as essentially the content of an appearance, by which they are contributed to by thought, rather than the content of an appearance that has no such contribution. By being a product rather than a precedent of thought, the self is thus cast as constructed. The self will be illusory insofar as it appears to underlie and precede thought, rather than to exist, as it actually does, as the content of thought.

James' account has significantly influenced those of three contemporary thinkers: Owen Flanagan (1992), Antonio Damasio (1999) and Daniel Dennett. Flanagan believes that it is a mistake to suppose that there is:

> … an 'I' that stands behind all conscious experience [and] … constitutes the core of the self, our conscious control centre, the source of all action plans, and the agent for whom all experiences accrue before being filed

> for future reference or discarded … the mind's 'I' is an illusion (Flanagan, 1992, 177) … The illusion is that there are two things: on one side, a self, an ego, an 'I,' that organizes experience, originates action, and accounts for our unchanging identity as persons and, on the other side, the stream of experience. If this view is misleading, what is the better view? The better view is that what there is, and all there is, is the stream of experience. 'Preposterous! What then does the thinking?' comes the response. The answer is that 'the thoughts themselves are the thinkers' (James, 1892, 83)[5] … We are egoless. This, of course, sounds crazy, so I have a fair amount of explaining and comforting to do (Flanagan, 1992, 178).

Flanagan (1992, 177, 178) builds his explanation of our sense of self on 'James' idea that the self is … an after-the-fact construction, not a before-the-fact condition for the possibility of experience'. Flanagan thus renders the self as constructed (and hence illusory) in the same general manner as James and Hume: the core of the illusion being pinned mainly on unity and identity of our 'conscious control centre' (viz., 'organizes experience … accounts for our unchanging identity as persons'). There is however one point, in connection with Flanagan, that should be clarified. When he mentions the illusory status of 'the mind's "I" ' – equivalent to the self – he sometimes gives the impression that this could be merely a philosopher's illusion, 'a tempting idea' which, with the help of precise thinking, can be eradicated (Flanagan, 1992, 177, 182–83). While there have indeed been many philosophers whose theories have given the self a serious unconstructed status that matches the content of our common conviction (often with extra bells and whistles, such as immortality) it should never be forgotten that this *common deep-seated conviction* is what spurred such philosophers on in the first place. Disabusing oneself of such a philosopher's illusion – if the self is indeed an illusion – would therefore not entail disabusing *oneself* (*qua* subject) of the illusion conveyed in *harbouring* the sense of the self, namely, the conviction that one really is an ontologically unique, unconstructed thinker of thoughts (etc.). It would rather entail replacing any incorrect philosophical account, one that accords with our common conviction, with a correct philosophical account of how it is that we – including the well-informed philosopher – come to have this deep-seated but *mistaken* conviction.

Damasio (1999), who has also acknowledged the influence of James, sees the problem of consciousness as being divided into two. The first problem, he says, concerns how we get a 'movie-in-the-brain' with its multi-modal sensory input culminating in familiar mental images that convey aspects of physical objects as well as the thoughts we have about them (Damasio, 1999, 9). The components of such images are known as 'qualia' – raw sensory qualities such as the whiteness or fragrance of a rose. Explaining how the physical brain could give rise to such qualities occupies a large field in the philosophy of consciousness. The second problem of consciousness,

according to Damasio (1999, 9), concerns how the brain 'engenders a sense of self in the act of knowing [about the world through such images]'.

> The solution for this second problem … requires the understanding of how the images of an object and of the complex matrix of relations, reactions, and plans related to it are sensed as the unmistakable mental property of an automatic owner who, for all intents and purposes, is an observer, a perceiver, a knower, a thinker, and a potential actor.
>
> (Damasio, 1999, 10–11)

Like Hume, James and Flanagan, Damasio does not think that this sense of being the observer and owner of the mental images is explained by there *really being* such a self that plays the role of owner and observer:

> This second problem is all the more intriguing since we can be certain that the solution traditionally proposed for it – a homunculus creature who is in charge of the knowing – is patently incorrect. There is no homunculus, either metaphysical or in the brain, sitting in the Cartesian theatre as an audience of one and waiting for objects to step into the light … In effect, the second problem is that of generating the *appearance* of an owner and observer for the movie *within the movie* ….
>
> (Damasio, 1999, 11)

In essence, Damasio holds that mental images from the somata-sensory modalities, hence, perspectivally ownable *objects* of consciousness, 'constitute knowing and sense of self' (Damasio, 1999, 128) and that 'the essence of core consciousness is the very thought of you – the very feeling of you – as an individual being involved in the process of knowing of your own existence and of the existence of others' (Damasio, 1999, 127). The very thought and feeling of oneself is conveyed in 'storytelling', 'from non-verbal imagetic to verbal literary' by which the brain, for adaptational reasons, 'map[s] what happens over time *inside* our organism, *around* our organism, *to* and *with* our organism, one thing followed by another thing, causing another thing, endlessly' (Damasio, 1999, 189). Our most basic sense of self, according to Damasio, is thus not produced by an actual, pre-existing self, but by images generated by the brain, which tell wordless, primordial stories about what happens to the organism as it goes about engaging with the world:

> Whether we like it or not, the human mind is constantly being split, like a house divided, between the part that stands for the known and the part that stands for the knower … The story contained in the images of core consciousness is not told by some clever homunculus. Nor is the story really told by *you* as a self because the core *you* is only born as the story is told, *within the story itself*. You exist as a mental being when primordial

> stories are being told, and only then; as long as primordial stories are being told, and only then. You are the music while the music lasts.
>
> (Damasio, 1999, 191)

Damasio regards that virtually *all* features of the self – knowing (witnessing), elusiveness, unity, identity, agency, (perspectival and personal) ownership and boundedness – to be entire mental constructs, in and of themselves. It is hence in quite an extensive manner that Damasio means 'the core you is only born as the story is told'. In the context of modern neurobiology (Damasio is first and foremost a neurologist), he elaborates upon certain features of James' account, in particular, how boundedness of self is constructed primarily by the emotions pertaining to individual concern (corresponding closely with the Buddhist notion of *taṇhā*). In a later chapter, I describe in more detail this aspect of Damasio's theory, since I believe that it provides independent support for the Buddhist contention of how boundedness lends the self its constructed status.

Let us finally consider Dennett. In his chapter 'The Reality of Selves' (in *Consciousness Explained*), Dennett wastes no time in dismissing the reality of self:

> Are there entities, either *in* our brains, or *over and above* our brains, that control our bodies, think our thoughts, make our decisions? Of course not! Such an idea is either empirical idiocy (James' 'pontifical neuron') or metaphysical claptrap (Ryle's 'ghost in the machine').
>
> (Dennett, 1991, 413)

Damasio has noted that Dennett's positive account of consciousness, against which the self is deemed an illusion, is one which depicts consciousness and the self as a 'post-language phenomenon' (Damasio, 1999, 188). Dennett thus conceives of the self in its broader capacity of being an 'owner of personality', a personal owner of a rich and detailed history that belongs exclusively to it. Like James, Dennett believes that the sense of such a self is grounded not in an actual self, but in a mental construct:

> But the strangest and most wonderful constructions in the whole animal world are the amazing, intricate constructions made by the primate, *Homo sapiens*. Each normal individual of this species makes a *self*. Out of its brain it spins a web of words and deeds and, like other creatures, it doesn't have to know what its doing; it just does it … Our tales are spun, but for the most part we don't spin them; they spin us. Our human consciousness, and our narrative selfhood, is their product, not their source (Dennett, 1999, 416) … These strings or streams of narrative issue forth *as if* from a single source … their effect on any audience is to encourage them to (try to) posit a unified agent whose

> words they are, about whom they are: in short, to posit a *centre of narrative gravity* (Dennett, 1999, 418).

It should be clear enough that on Dennett's account, as with the preceding ones, a self, *qua* underlying agent that unifies our experience and creates our thoughts, is both constructed and illusory.

3. Buddhist and the Western accounts of 'no-self': Summarising the similarity and differences

We have alluded to a number of well-known Western thinkers with accounts that converge with that of Buddhism insofar as the self is regarded, as a whole, to be a construct and hence (by our definitions) an illusion. We are also able to see how the accounts diverge, namely, through the manner in which the self is considered constructed. The Western accounts diverge from the Buddhist account insofar as they regard many ascribed features – viewed as inherently unconstructed by Buddhism – to be entirely constructed. Unity, unbrokenness and invariability (often embedded in talk of personal identity) are most commonly alluded to in this capacity. Buddhism regards these features to be constructed only to the extent that their impression is distorted through mistaken ascription (on behalf of witnessing) to a bounded, personal owner. For example, when the unbrokenness and invariability native to witnessing become welded to the impression of a bounded self, they yield the impression of not only immediate, but longer-term identity. This longer-term identity involves distortion not intrinsic to the witnessing and *to this extent*, identity (like the self as a whole) is a mental construct. Yet its core aspect of unbroken invariability (cognised from one conscious moment to the next) is owed to what witnessing brings to the self-sense, and this is *not* considered a mental construct. Unlike on Hume's account, the impression of identity – indeed the sense of self – is not thoroughly rooted in an ontology of impermanence. Our analysis should therefore enable us to see more clearly how two parties (Buddhist and Humean) can deny the overall reality of self while varying significantly in their details of *how* the self lacks reality.

In light of this analysis, it will be useful to reiterate the analogy that was used in the Introduction to help cognize the difference between the Buddhist and the typically Western account of no-self. Suppose that two people dream about a shrill voice. The shrill voice is in both cases a construct, the content of an appearance contributed to by thoughts, images and so forth. We can suppose that the first dream is woven around the sound of an alarm clock, which lends <the shrill voice>, as the content of the appearance, its piercing quality. To the extent that the shrillness is contributed to by the dream-independent alarm sound, the shrillness is not, in itself, a mental construct. It is only a construct to the extent that it is ascribed, in the dream, to the shrill voice and is distorted through this assumption. We can suppose that in

the second dream, the sound of the shrill voice is not woven around any alarm but is dreamt up entirely. The shrillness ascribed to the voice, as well as the shrill voice itself, is a mental construct. So <the shrill voice>, when contributed to by the alarm sound, is analogous to <the self> as construed on the Buddhist account, with many of its contributing features inherently unconstructed. <The shrill voice> that is entirely dreamt up is analogous to <the self> as construed on the Western accounts (especially Damasio's), with many, if not all of the key features entirely constructed.

Conclusion

In this chapter we articulated a way in which parties from different metaphysical backgrounds can uniformly consider the self to lack reality. The self, if it lacks reality, would lack reality by being illusory, that is, by purporting to exist in a manner in which it does not actually exist. The self purports to exist in a manner that is inherently unconstructed by any thoughts, perceptions or objects of consciousness that can be perspectivally owned by it. That is, it purports to be that thing which precedes and thinks the thoughts, owns the experiences, always standing apart from them in this fashion. If illusory, then the self would exist in a manner that is constructed. It would amount to nothing but the content of an appearance that is contributed to by those very objects that it purports to stand apart from. The thoughts will be helping to think the self rather than the self thinking the thoughts.

Within this converging framework of self-as-construct-and-illusion, we looked at differences in the way that the self can be deemed a construct (and hence illusion). The Western accounts considered in this chapter – those of Hume, James, Dennett, Flanagan and Damasio – all showed evidence of regarding the self as more thoroughly constructed than in Buddhism. In particular, the Western thinkers regarded the individual features of unity, unbrokenness and invariability (short-term and long) to be thorough constructs. The Buddhist account of self, as we know, considers *boundedness* (as brought about through a sense of personal ownership, agency and self-identification) to be the primary source of the self's illusory status. To the extent that their impression is not distorted by that of boundedness, the (misascribed) features of witnessing, unity, elusiveness, unbrokenness and invariability are considered, in and of themselves, to be unconstructed.

It is a primary goal of this project to demonstrate that the Buddhist account of no-self is correct, with a metaphysically neutral 'witnessing' referred to in lieu of nibbānic consciousness. In the following chapter I begin to advance this position by arguing that the notion of witnessing must be analysed in such a way that implicates those features considered unconstructed in Buddhism. This will make it more likely that witnessing, as advanced on the Buddhist account, will 'carry' these features – themselves unconstructed – into an overall construct of the self.

6
Linking Problems of Consciousness with Awareness

Introduction

On the Buddhist account, features of unity, unbrokenness, invariability (and by implication, elusiveness) are 'imported' into the sense of self through non-illusory (and pre-nibbānic) witness-consciousness, to be warped – by *taṇhā* and identification – into the impression of a bounded personal owner that has those features. If this Buddhist analysis is correct, then upon examining the relation between features ascribed to the self, we should expect to find a special link existing between witness-consciousness and its alleged 'characterising aspects' of unity and so forth – a link that does not extend to boundedness, personal ownership or agency. The purpose of this chapter is to argue that there is indeed a special link existing between witness-consciousness and those aforementioned features, when witness-consciousness is construed as having an intrinsic subjective character (something the self also purports to have). As the proposed link will be conceptual in nature, there will be no need, at this stage of the argument, to establish that such features actually exist. I will thus argue that it is part and parcel of the very concept of such 'felt' witness-consciousness – or *awareness*, as I shall call it – that it be analysed with reference to the concepts of unity, elusiveness, unbrokenness, invariability and unconstructedness – *as these features present in ordinary conscious states*. The concepts of these features – *as they present in ordinary conscious states* – will in turn be analysed with reference to the concept of awareness.

The emphasis on 'as the features present in ordinary conscious states' is an important one. Just as it is not being claimed (in this chapter) that awareness and those features really exist, it is not being claimed that if awareness were to exist, then it would be *objectively* unified, unbroken, invariable, etc. – at least as those notions are often construed in philosophical literature. From the outset, I have been careful to define those features in a way that makes it plausible for them to be ascribed to the ordinary self – and this has put notable constraints on how they are to be understood. For example, the kind of (felt)

permanence ascribed to the (Buddhist notion of) self in Chapter 3 had to be one that resisted an implausible reading that immortalised the self or disallowed for natural breaks in conscious life such as deep sleep. The notion of 'unbrokenness' referred to in this book is supposed to capture no more than just this aspect – namely, that aspect by which, from the first-person perspective, there is something about one's conscious life that (from one conscious moment to the next) is unbroken in its presence. (The fact that we ascribe such a feature to ourselves was established in Chapter 4.) So when it is argued in this chapter that the concept of awareness implies the concept of unbrokenness (and vice versa), it is not being claimed that awareness (if it existed) would be unbroken in a blatantly objective manner that defies deep sleep and death. Rather it is claimed that if awareness were to exist, then from the first-person perspective it would have to *present*, in a given stretch of conscious time, *as unbroken* rather than gappy, and (conversely), that where there is a *sense of unbrokenness* in one's conscious state, from the first-person perspective, such unbrokenness must *present* as qualifying awareness. Note that when unbrokenness is defined with reference to the first-person perspective, there is no discernable difference between '*presents* as unbroken' and '*is* unbroken'. *Mutatis mutandis* for the features of invariability and so forth.

Despite what may seem to be a rather conservative reading of the key features being linked with awareness, those features, construed as such, have generated longstanding puzzles of subjectivity in the philosophy of mind. As a corollary to this chapter's main claim, I will suggest that because it singularly captures so many puzzling features, the concept of awareness – barely acknowledged in Western philosophy – is an essential component of phenomenal (subjective) consciousness. The concept of awareness should thus be addressed in any theory of consciousness that purports to completely explain its subjective character.

I will then argue that the link between the concept of awareness and those of its features, does not extend to the concept of boundedness (central to the notion of self), which depicts that feature of ontological uniqueness as borne out through the role of personal owner. While boundedness must be analysed with reference to awareness (simply because the self, as a bounded personal owner, will be also a witnessing subject that is unified and so forth), the reverse implication will not hold. Awareness will not implicate boundedness, demonstrating that a conceptual link will hold *only* between those concepts of the self's features that Buddhism ascribes to (pre-nibbānic) witnessing. Such a scenario will increase the likelihood that *boundedness* underlies the self's constructed status – if the self is indeed constructed – a Buddhist contention to be fully argued for in Chapters 8 and 9. These later chapters will also discuss the happiness-seeking (and *dukkhā*-avoiding) urge in its relation to boundedness and *taṇhā* – although there will not be the space to determine its independent ontological status (remembering that Buddhism grounds a measure of this feature in *nibbāna*).

We must re-iterate that this chapter's task is *not* to demonstrate that awareness and its specifying features of unity, etc. have the unconstructed reality that they purport to have, and would *have* to have if the Buddhist account was correct. The task of demonstrating the independent reality of awareness and so forth is postponed until Chapter 7. This current chapter's discussion will remain neutral on the ontology of awareness and associated features.

1. Awareness as a concept of consciousness

Before we can define awareness through the features of unity, elusiveness and so forth – and these features in terms of awareness – the core concept of awareness must itself be clarified. The term 'awareness' will henceforth be used in this book to denote a witness-consciousness with an *intrinsic phenomenal character*, a subjective character of its own that it brings to all conscious experience. This phenomenal dimension, although arguably implicit in non-philosophical usages of 'awareness', immediately contrasts the term with a more technical usage that is sometimes apparent in philosophical circles. In such instances, 'awareness' refers to specific *functions* that consciousness may have (such as its disposition to cause patterns of behaviour) – with any phenomenal dimension being ignored or denied. For example, David Chalmers (1996a, 28–31) uses the term 'awareness' to refer to the *non-conscious* functional analog of phenomenal consciousness. In the context of this book, however, the term 'awareness' will be used to refer, more in line with everyday usage, to witnessing with phenomenal character. There will be something it is like to *be aware simpliciter*.

The core concept of awareness is hence as follows. Awareness in this sense:

1. involves the 'subject' side of the apparent subject/object dichotomy, specifically
2. the *modus operandi* of witnessing and
3. has an intrinsic phenomenal character that is owed to its subjective status rather than to any objects – attended or peripheral – that the subject might be conscious of.

1.1. More on the *modus operandi* of witnessing

I have already spoken at some length about witnessing (in earlier chapters), so I include here just a brief review. Witnessing (or witness-consciousness) is to be understood as the most generic mode of conscious apprehension, involved in, but not identical to such specific modes as thinking, sensing, perceiving, introspecting and remembering. It is the essence of knowing that is common to all these modes; it is that by which we can know we are thinking, sensing, perceiving, introspecting or remembering. Witnessing

can be attentive or inattentive. Consider what happens when one attends to an object. The object is the focus of attention, so what we call 'witnessing' is directed mostly towards the object. Nevertheless we can argue that there is inattentive witnessing of such objects as background noise and proprioceptive feelings. Evidence of inattentive witnessing is one's ability to recollect being dimly aware of these items all along. The fact that it makes sense to speak of inattentive witnessing or to say 'there was witnessing of the noise though I did not attend to it' suggests witnessing to be more basic than attention. A useful analogy for the relation between witnessing and attention is that of a flashlight, whose beam can be concentrated and focused on various objects, while nevertheless casting a dimmer pool of light on objects not focused on. Witnessing can be likened to the light in general, while attention can be compared to the focusing or directing of the beam on this object and that. Even with the beam focused, there is a surrounding, dimmer circle of light, illuminating objects not focused on, analogous to the inattentive witnessing.

1.2. The intrinsic phenomenal character of awareness

Awareness is defined as witnessing that is intrinsically experiential, imparting, on the metaphor used by Eastern 'guru' Ramana Maharshi (Osborne 1971, 133), a common 'flavour' to all conscious states. Barry Dainton (2002, 32) refers to it as 'tangible' (as opposed to 'pure' which lacks any phenomenal character).[1] The important matter is that its phenomenal character be construed as intrinsic to awareness itself. This means that its phenomenology – or 'what-it-is-like-ness' – is not reducible to the spectrum of experiential qualities that, whether attended to or peripheral, characterise the sense-differentiated objects of experience, including the sensory modes through which objects are perceived. It is thus not reducible to the sensory qualities specifically associated with the perceptual modalities of vision, sound, taste, smell, touch, proprioception and with cognitive deliverances of emotion, mood, conceptual thought, memory or imagination. There is nevertheless, on the proposed concept, something it is like to be aware simpliciter and this quality of awareness (or knowing) imparts its own generic 'flavour' to all conscious states. A useful analogy may be that of a beam of white light being diffracted by a prism into a spectrum of colour. While each colour (sensory-experience) will differ from the others, all will share the unifying and generic character of luminosity (awareness).

Now since the language of phenomenology is generally devoted to describing objects of experience, it is naturally difficult to convey a phenomenal, subjective character that does not pertain to objects. While such difficulty has, I believe, deterred many thinkers from recognising the proposed concept of awareness, others have attempted to characterise what they hold to be a 'background' aspect of human experience, an aspect that seems to reduce not merely to the observed objects. From the Eastern

tradition of Advaita Vedanta, Nisargadatta Maharaj (1980) refers to this aspect of experience as 'conscious presence' (Powell, 1994, 87), while Arthur Deikman (1996, 350), from the Western tradition, describes it as 'the subjective sense of our existence'. C.O. Evans (1970, 150) talks about a 'lively sense of presence' that is integral to our sense of self while Chalmers (1996, 10) alludes to a 'deep and intangible' 'phenomenology of self', which he likens to a 'background hum'. I propose that we call this background phenomenology 'the subjective sense of presence'. I prefer 'presence' to 'existence' because 'presence' conveys the double meaning of being present with respect to space (as in the perpetual here) and present with respect to time (as in the perpetual now) – and we arguably experience a subjective sense of presence through both the *here* and the *now*.

With regard to time, especially, our sense of the present moment seems to emanate either from or through the subject, rather than seeming attributable to objects of which the subject is aware. That is to say, our sense of the present moment does not seem to originate from objects of consciousness; rather, objects (such as perceptions or thoughts) seem to occur to a subject in the present moment. While this point is speculative, the sense of present-momentness is perhaps the most immediate way to discern what we are calling the 'phenomenal character' of awareness, and the most obvious clue that there could be more to conscious experience than just object experience.

Our core concept of awareness thus conveys a type of experience – a subjective sense of presence – that is unmediated by any specific quality pertaining to objects, outer or inner. In our concept, this subjective sense of presence intrinsically qualifies the witness-consciousness that is the *modus operandi* of a subject.

This completes our outline of the core concept of awareness as (a) involving the subject side of the apparent subject/object dichotomy such that it is (b) identical to the subject's *modus operandi* of witnessing and (c) of an intrinsic phenomenal character. While a handful of current Western thinkers (often familiar with Eastern traditions) give attention to a version of this concept (or something similar): for example, Ken Wilber (2001), Arthur Deikman (1996), Robert K.C. Forman (1998), Jonathan Shear (1996, 1999) and David Woodruff Smith (1986), 'awareness', as portrayed here, is not a mainstream concept in contemporary Western philosophy of mind (although historically, thinkers such as Moore and Hamilton may have endorsed versions of it).

Before moving on, I should relate this concept of awareness to a somewhat similar notion endorsed by thinkers such as Victor Caston (2002), Uriah Kriegel (2004) and Greg Janzen (forthcoming). Following in the tradition of Brentano and Aristotle, they allude to a concept of consciousness that, despite its similarity, may harbour an important difference. The Aristotelian/Brentanian version builds object-directedness into the concept of consciousness, such that a person's witness-consciousness (it would seem)

gets its elusive phenomenal character from reflexively taking itself as a secondary 'object' (the primary object being the attendable target of thought or perception). The question, of course, is whether this 'on-the-side secondary object' is something that (a) can be attended to in principle, hence conforming to the definition of object outlined in this book, or (b) better fits what I refer to in Chapter 5 as 'content' (having an 'aboutness'). Kriegel (2004, 194) claims that the on-the-side 'object' can (like the primary object) be attended to, but it is not clear that his version of the theory would be supported by all those who develop an Aristotelian or Brentanian theory of consciousness.[2] The concept of awareness developed in this chapter could be compatible with an Aristotelian/Brentanian theory if the latter was interpreted in line with (b) rather than (a). For I will argue that through its conceptual link with such features as unity and so forth, awareness will have a content that effectively renders it to be *about* itself.

A further general difference is that the Aristotelian/Brentanian notion of consciousness is construed as intrinsically object-oriented in having to have at least one attendable object in its purview (the primary object). The concept of awareness (as with *witnessing* in Chapter 1) is left uncommitted on this front. It could turn out that awareness must, as with the Aristotelian/Brentanian consciousness, occur in a conscious state with an attendable (primary) object (and so be part of a more complex conscious state), or it could turn out that awareness, as in Buddhist analysis, could exist without any attendable objects at all. I do not try to adjudicate this issue.

2. Linking problems of consciousness with 'awareness'

In this section it will be argued that several classic puzzles of phenomenal (subjective) consciousness, featuring elusiveness, synchronic unity, unbrokenness, and invariability cannot be specified without recourse to the concept of awareness and that the concept of awareness cannot, in turn, be specified without reference to these features. Through its implication in these classic puzzles, awareness will also be shown to be an essential component of phenomenal consciousness. (Hence, should phenomenal consciousness exist, then, so will awareness.)

The strategy throughout this section will be to first outline each of the features as they (or most of them) occur in the context of traditional puzzles for subjective consciousness. In doing this, I will be staying neutral on whether subjectivity, consciousness or the notion of a subject is to be defined with reference to awareness (or witnessing), unless the puzzle specifically alludes to consciousness in that manner. The next stage will be to demonstrate that the concept of awareness does indeed best fit the notion of consciousness that makes each feature puzzling. That the concept of each such feature must be specified with reference to the concept of awareness

(and awareness in terms of these features) will aptly demonstrate a conceptual link that exists between awareness and those features.

2.1. Elusiveness

In Chapter 4, I argued that elusiveness is a feature ascribed to the ordinary self; in this section it is argued that elusiveness should be specified with reference to awareness in particular. In a supporting passage, I quoted Roderick Chisholm (1969) where he drew attention to the centrality of elusiveness as a problem of consciousness, framing the problem as one about subjectivity in particular. Here is the passage in full:

> The two great traditions of contemporary western philosophy – 'phenomenology' and 'logical analysis' – seem to meet, unfortunately, at the extremes. The point of contact is the thesis according to which one is never aware of a subject of experience [… when] as Hume put it, we enter most intimately into what we call ourselves. Thus Sartre seems to say that, although we may apprehend things that are pour-soi, things that are manifested or presented to the self, we cannot apprehend the self to which, or to whom, they are manifested – we cannot apprehend the self as it is in itself, as it is en-soi. And Russell has frequently said that the self or subject is not 'empirically discoverable'; Carnap expressed what I take to be the same view by saying that 'the given is subjectless'.
>
> (Chisholm, 1969, 94)[3]

To recapitulate, 'elusiveness' pertains to the fact that one is never aware of a subject of experience, or more accurately, to the fact that the subject cannot observe itself as an observing subject that is simultaneously its own object of experience. The subject is systematically elusive to its own attentive purview. On the face of it, we might wonder why there is a puzzle at all. Elusiveness is a ubiquitous phenomenon, not confined to subjects of experience, and most cases of it are innocuous. For example, no metaphysical enquiry has been launched into why a human eye cannot directly see itself, or why one cannot jump on the head of their own shadow. Such facts are a matter of logic, and philosophers such as Gilbert Ryle (1966) have argued that people are confused if they think that there is more to 'the elusive subject' than this.[4] Just as a video camera cannot, as a matter of logic, directly observe (that is, video) itself, the subject, as the locus of an observing perspective, cannot directly observe itself.

Even if Ryle is ultimately correct, the fact of the matter is that philosophers from 'the two great traditions' do think that there is more to 'the elusive subject' than a mere quirk of logic. We must ask: why do they think that? Among the panoply of elusive phenomena, why is the subject singled out as the puzzling case? The answer seems to be that while on one hand, we as subjects are systematically eluded from our own direct observation, on

the other hand we have some intuition – a conscious sense or feeling – that we do, simultaneously and immediately, experience ourselves as observing subjects. This element of experience must pertain (or seem to pertain) to the subject's *modus operandi* if the case of 'elusive subject' is to be distinguished from the vast stock of innocuous elusive phenomena (such as the video camera). What is puzzling, then, is that something about this *modus operandi* suggests an immediate experiential sense of our subjective existence, and yet this subjective existence can never be attended to *in* our experience. The subject's *modus operandi* thus seems to involve a phenomenological tension. Here is a formulation of the elusiveness phenomenon, in light of the puzzling feature just outlined:

> Elusiveness of the subject of consciousness (through its modus operandi): On the one hand, the subject is elusive to itself. That is, it cannot attentively observe itself. On the other hand, the subject has an immediate experiential sense of its own existence. Hence, the puzzle. The subject – in virtue of its modus operandi – seems to be that which is both elusive to attention and yet experienced.

Now this book has already defined the subject's *modus operandi* as witnessing. For the current purposes, however, it would be presupposing too much to define the subject's *modus operandi* as witnessing. We want, after all, to demonstrate that a kind of witnessing, in particular, an intrinsically phenomenal witnessing (viz., awareness), is indeed the *modus operandi* of subjecthood referred to in the puzzle of elusiveness. So we must answer the question: what concept of consciousness best captures the subject's *modus operandi* by seeming both systematically elusive to attention and yet experienced? In both this and the following section I first investigate whether popular, alternative concepts of consciousness depict what is puzzling about the case at hand. One concept will be popular although unlikely, while the other concept(s) may initially seem to fit the bill and be a close rival to that of awareness. I then argue that the concept of awareness is the only concept that truly captures the puzzling feature.

First consider 'phenomenal object consciousness'.[5] This notion equates consciousness with its phenomenal objects: specifically, qualia (raw feels) that pertain to the five senses and cognition (including specific feelings pertaining to thoughts and emotions). On this notion, a moment of conscious experience will be exhausted by whatever collection of phenomenal objects one is conscious of at a given time, for example: redness, dizziness, sadness and the feel of a chair.

Phenomenal object consciousness can be easily dismissed as a candidate for the elusiveness puzzle. While such qualia are puzzling in their own right, the puzzle is not that they seem to systematically elude attention. Indeed, it is because such qualia are so typically unelusive that theorists are able to

ponder such questions as why red and green seem complementary colours, or how the physical brain can give rise to such a variety of novel qualities. Elusiveness pertains rather to the subject of these phenomenal objects, a subject whose *modus operandi* seems to escape the purview of normal, in-the-face observation. 'Phenomenal object consciousness' is hence a non-starter, involving the wrong side of the subject/object dichotomy – to the objects apprehended rather than to the apprehending subject. It cannot therefore depict the *modus operandi* of 'subject' in our articulation of the elusiveness puzzle.

Consider now 'higher-order' theories of consciousness. On typical versions of these theories, a mental state is conscious by virtue of the fact that it is the content of a higher-order mental state, a state to which the content is presented. This higher-order state, be it a thought or perception, confers consciousness upon its directed content without actually being intrinsically conscious of itself (Caston, 2002, 753–755). The higher-order state would satisfy the elusiveness requirement, since its inevitable directedness at other mental states would mean that it could not directly observe itself (if conscious, it is always in virtue of being the target of another higher-order state). However, the higher-order state clearly does not satisfy the requirements for the elusiveness puzzle, since there would seem to be nothing it is like to be in such a state *per se*. Since it *ex hypothesi* seems to lack any intrinsic phenomenal character, a higher-order mental state could thus add none of its own such character to any given conscious experience. The phenomenal dimension to a given conscious state will be exhausted by those observable objects to which the higher-order state is directed. Nor would the *phenomenal objects* satisfy the requirements of the elusiveness puzzle: while experienced, they would not be elusive to observation. Hence, 'higher-order consciousness' – whether construed with emphasis on the 'content-made-conscious' or on the higher-order 'consciousness-conferring mental state' – cannot be that aspect of the subject that seems, simultaneously, to be both unobservable and yet experienced.

Let us now consider a more serious contender: that of 'unprojected (or peripheral) consciousness', proposed as a candidate for 'the subject' (its designated *modus operandi*) by Evans (1970). Evans (1970, 104) writes: 'I give the name 'unprojected consciousness' to those elements of consciousness that together make up the background of consciousness when attention is paid to an object'. Basically, Evans (1970, 104–107) holds that our field of consciousness is polarised into two areas: that of the objects being attended to – 'the foreground' – and that to which we are not attending but are nevertheless consciously aware – 'the background', or, as he calls it, 'unprojected consciousness'. Objects of attention occur, and must occur, he says, against a background of unprojected consciousness that is composed of such elements as thoughts and feelings. Any mental contents which happen to constitute a moment of unprojected consciousness need

not be there the next moment: they are constantly shifting, some becoming objects of attention, others ceasing to be elements of consciousness. What is important is that although the contents are mobile, the twofold structure of consciousness into foreground and background remains fixed. Our awareness of elements within unprojected consciousness is of course dim. At the time of experience, we do not individually differentiate the elements of unprojected consciousness; otherwise we would, by definition, be attending to them. Rather, he claims, the experience is of an undifferentiated whole which lends phenomenal character to our consciousness, but can be discerned as separate elements only in retrospect, for example:

> … when eventually the pain in the blistered heel is noticed its presence need not come as a complete surprise. The person may recollect that he had after all been dimly aware of it all along. It is this form of awareness to which I am referring when I claim that the elements of unprojected consciousness are elements of awareness. If they were not, there would be no justification in including them in consciousness – unprojected or otherwise.
>
> (Evans, 1970, 107)

Now it makes no sense, says Evans, to attend to elements in unprojected consciousness. For the moment we shift our attention to elements in the background of awareness, they become, by definition, objects of attention in the foreground, while a new lot of elements form the background. For this reason, he says that unprojected consciousness, while experiential, is systematically elusive to attention (Evans, 1970, 148–149). Since he regards a defining feature of 'the subject' (which he refers to as 'the self') to be its elusive yet experiential nature – and so in accordance with the above specification of the elusiveness puzzle – he sees it as a natural step to identify that aspect of subject-self with unprojected consciousness:

> There is a parallel with the 'I' which, as Ryle has argued, is systematically elusive too in the sense that the 'I' cannot be objectified by attention (Evans, 1970, 149) … I am suggesting that it is indeed true that the subject-self can never itself become an object of experience, and I am maintaining that this has nothing to do with the nature of the self – transcendental or empirical. It is put down to nothing other than the way attention operates. This enables it to be asserted in all consistency both that the self as subject is experiential, and that it is never presented as an object of experience. Furthermore, it obviates the necessity of treating the self as something unknown in itself. The theory overcomes the paradox that the self, although discoverable in experience, is never an object of experience, and in the process removes the main prop holding up The Pure Ego Theory of the Self. The essence of the matter is, on my view, that

> the self is experiential (i.e. is composed of elements of consciousness), but is never known as an object of experience. This is one of the factors that accounts for the view that the self lies behind its experiences. It also explains why we have such a lively sense of the presence of the self, and why we are so nonplussed by denials of the self's existence … (Evans, 1970, 150).

At first glance, the elements that at any one time form an unprojected consciousness nicely fit the specification for the elusiveness puzzle. The elements are elusive, in that attention can never be focused upon them without them ceasing to be elements of unprojected consciousness. At the same time they lend to our conscious life a rich background phenomenology, fitting the bill of being experiential as well as elusive. On closer inspection, however, the concept of unprojected consciousness is unsuitable for specifying the elusiveness puzzle – at least on the level of how things appear, the level that concerns us here. Subjects apprehend, observe, notice, attend, perceive and introspect. That is, the subject must carry a disposition for apprehending objects, whatever the mode of apprehension. The problem with identifying the subject's *modus operandi* with unprojected consciousness is that the latter does not fit this bill, since its elements do not seem, under any stretch of the imagination, to apprehend objects. It does not remotely seem, for instance, as if a peripherally registered pain in the heel, along with the rest of unattended-to elements, is involved in attending to the objects of attention, such as a speech. And yet that is how it should seem, if unprojected consciousness is to substitute the *modus operandi* of 'subject' in the elusiveness puzzle. Rather, it seems as if a subject – one and the same subject – is, via its *modus operandi*, inattentively noticing the pain in the heel, etc., while attending to a speech. This is the subject whose *modus operandi* seems problematically elusive.

On the level of appearances then, the elements of unprojected consciousness are far more suitably identified with observable objects of consciousness than with the elusive *modus operandi* of an observing subject. Indeed, any item which may at one moment form an element in unprojected consciousness, such as a pain in the heel, can be attended to the very next second – which clearly renders the item an object by this book's definition. And even while concurrently elusive to attention, the elements of unprojected consciousness are not elusive to broader modes of apprehension such as inattentive noticing – modes of apprehension that seem, in and of themselves, to be elusive to their own observation. So while Evans's concept of unprojected consciousness does capture something elusive and experiential, it is ill suited to convey the subject's *modus operandi* at the heart of the elusiveness puzzle. We must look elsewhere for a suitable concept of consciousness.[6]

Let us now consider whether our concept of awareness can convey the kind of consciousness that is implicated in the subject's puzzlingly elusive

yet experiential *modus operandi*. We can immediately note that awareness, through its component of witnessing, involves the broadest mode of conscious apprehension, whether attentive or inattentive. So, unlike unprojected consciousness, awareness is in the right ballpark to fit what is required by the subject's *modus operandi*. Awareness also satisfies the requirements of the elusiveness puzzle, since its very definition implies a phenomenal character – a subjective sense of presence – that is not owed to any object of experience. Since an object is by definition anything that can be attended to (by a subject), it follows that the phenomenal character of awareness, not owing its character to objects, cannot be attended to by a subject. Since awareness cannot be attended to by a subject, it is elusive in the required sense. Denoting something both experiential and yet elusive to direct observation, the concept of awareness naturally fits our specification of the elusiveness puzzle. Unlike the concept of unprojected consciousness, which also denotes something elusive and experiential, the concept of awareness captures what is distinctive about the subject's *modus operandi*.

The concept of elusiveness (featuring in the puzzle) hence implies the concept of awareness, and the concept of awareness, in turn, implies the concept of elusiveness.

2.2. The synchronic unity of consciousness

We now turn to the 'synchronic unity of consciousness' – also enlisted, in Chapter 4, as an assumed attribute of the self. In this section, I attempt to show that the concept of synchronic unity is tied up, in particular, with the concept of awareness. Bayne and Chalmers have articulated the phenomenon of unity in the following way:

> At any given time, a subject has a multiplicity of conscious experiences. A subject might simultaneously have visual experiences of a red book and a green tree, auditory experiences of birds singing, bodily sensations of a faint hunger and a sharp pain in the shoulder, the emotional experience of a certain melancholy, while having a stream of conscious thoughts about the nature of reality. These experiences are distinct from each other: a subject could experience the red book without the singing birds, and could experience the singing birds without the red book. But at the same time, the experiences seem to be tied together in a deep way. They seem to be unified, by being aspects of a single encompassing state of consciousness. (2003, 23)

From this, we can formulate the phenomenon that underlies the puzzle of synchronic unity:

> Synchronic unity of consciousness: Whenever a multiplicity of objects appear to a subject at any one time, they seem phenomenally unified

> to the subject by being aspects of a *single encompassing state of consciousness*.

The puzzle is how normal experience, so obviously characterised by a multiplicity of objects, can at the same time seem phenomenally unified to a subject. Which concept of consciousness best captures the phrase 'single encompassing state of consciousness' and hence best conveys the conscious unity that, from the first-person perspective, seems to belie the obvious complexity? As with elusiveness, we will countenance a few possibilities, beginning with the least plausible: 'phenomenal object consciousness'.

For similar reasons as before, talk of 'phenomenal object consciousness' is a poor substitute for the phrase 'a single encompassing state of consciousness'. Basically, the conveyed experiences lie on the wrong side of the subject/object divide. Construed as objects, those experiences will appear as the complexities that are unified to a subject in a single state of consciousness, not as the single state of consciousness in which the complexities seem unified. Indeed, the encompassing state of consciousness does not appear as an object of consciousness at all. If it did appear that way, then it is likely that Hume, seeking the source of unity he termed 'simplicity', would have located it amongst his disparate perceptions. The concept of phenomenal object consciousness therefore fails to capture that feature which underpins the puzzle of synchronic unity.

Talk of 'higher-order consciousness' is also unsuited to capture what is meant by the phrase 'single encompassing state of consciousness'. The unity of consciousness, like elusiveness, is puzzling because it presents as something whose presence is intuited subjectively and pre-theoretically, from the first-person perspective. As with elusiveness, there is a phenomenological tension; the unity of consciousness subjectively seems to belie the complexity of its attentively observable content (the philosophical challenge being to explain this puzzling appearance). One does not have to be a theoretician or a philosopher to discern, in a given conscious state, a sense of unity underlying the multiplicity; unity does not seem like the mere abstract conclusion of a logical inference. If it did seem like a mere abstract conclusion, then unity would have never been puzzling in the way that it is. Now a higher-order mental state may well be unifying, but there is by definition no phenomenal character to this unity: the experiential dimension to such consciousness being exhausted by the multiplicity of objects to which the state is directed. The concept of higher-order consciousness (in virtue of the higher-order mental state) is therefore unsuited to capture that notion of consciousness implicated in the unity puzzle.

Another possibility lies in talking in terms of 'access consciousness', as the notion is defined by Ned Block.[7] We can quickly dismiss talk of 'access consciousness' as suitable for unpacking what is meant by consciousness in the unity puzzle. Like higher-order consciousness, access consciousness is not

defined with reference to any intrinsic phenomenal state. Rather, it is defined with reference to causal and functional roles in the cognitive economy. For a subject to be access-conscious of an object is for that object, as the content of a mental state, to be available for verbal report, reasoning and voluntary behaviour. For example, if I am access-conscious of a book, then that entails I can say things like 'there is a book' and do such things as pick it up and return it to the library. Since there need not be anything it is like for me to execute this exercise, any unity bestowed by access consciousness will not, by simple virtue of its being bestowed by access consciousness, be phenomenally felt.

We will now consider a concept of consciousness that Bayne and Chalmers (2003) think accurately conveys what seems unified about consciousness. Unlike the concepts expressed by the terms 'access consciousness' and 'higher-order consciousness', the proposed concept is of something inherently experiential, involving what Ned Block refers to as 'phenomenal consciousness'. Bayne and Chalmers (2003, 28) write that a mental state is phenomenally conscious 'when there is something it is like to be in that state … being in that state involves some sort of subjective experience'. Importantly, it would seem that this 'subjective experience' is not restricted to (attendable) objects of consciousness – it involves any dimension to consciousness that is experiential. Should there turn out to be phenomenal character inherent to the very subject of experience (through its *modus operandi*), then the concept of phenomenal consciousness will also cover that. This immediately puts the concept of phenomenal consciousness at an advantage over that of phenomenal object consciousness in its potential to capture what could be meant by the term 'single unifying state of consciousness'.

Bayne and Chalmers (2003, 26–27) continue their analysis of unity by referring to what they call 'subsumption'. A subsumptive state of consciousness – whether access or phenomenal – is a state which, at any given moment, subsumes or 'umbrellas' all the disparate 'object' experiences: auditory, visual, emotional, proprioceptive, etc. Subsumption is non-trivial because the subsuming state is not merely the conjunction of all these experiences, but is a single conscious state in its own right (Bayne and Chalmers, 2003, 27). It may, they say, be thought of as the 'subject's conscious field' which serves to unify the disparate elements, or as 'the singularity behind the multiplicity' (Bayne and Chalmers, 2003, 27). Given that access consciousness fails to properly capture what seems unifying about consciousness, a natural step is to analyse the subsumptive relation in terms of phenomenal consciousness. A subsumptive, phenomenally conscious state will thus be a state of consciousness in which there is something it is like for a subject to be in a single encompassing state that unifies the different elements. In the words of the authors:

> A set of conscious states is phenomenally unified if there is something it is like for a subject to have all the members of the set at once, and if this

> phenomenology subsumes the phenomenology of the individual states. (2003, 32) … A phenomenal state A subsumes phenomenal state B when what it is like to have A and B simultaneously is the same as what it is like to have A. (2003, 41)

Put another way, a set of conscious states will be phenomenally unified (at a time) when there is something it is like to be in a single subsuming state of consciousness that 'umbrellas' all the individual states. Can we be more specific about the phenomenology of the subsumptive state itself? We know that the subsumptive relation requires that its unifying state be not reducible to merely the conjunction of objects (e.g., sights, sounds, thoughts) that are subsumed by the state. We also know that this unifying state, on the proposed account, must involve a phenomenal character that is discernably separate from that of its objects (even if a given state of phenomenal consciousness must inseparably involve both subsumptive state and objects subsumed). We might therefore ask: from where could the phenomenal character of this subsuming state appear to emanate, if not merely from characteristics of objects that are subsumed by it? The answer must be: from that quarter involving the apparent subject of the conscious state. And the phenomenal character pertaining to the subject is strongly suggestive of awareness as its *modus operandi*. Hence, we have good *prima facie* reason to suppose that *awareness* is actually the concept of consciousness being appealed to by Bayne and Chalmers when, in an attempt to capture the unity of consciousness, they talk about a 'single encompassing state of consciousness'.

Before establishing this inference, we have to determine that the concept of awareness does indeed imply that of synchronic unity. Does our concept of awareness denote something that intrinsically yields a felt singularity that would appear to belie any observed multiplicity? The question is best approached by asking whether awareness could seem disunified at a time, in any conscious state. On reflection it seems not. Suppose awareness – viz., phenomenal witnessing – were to appear disunified to a conscious perspective at a single time. For a disunified X to be cognised as a disunity, in a single conscious moment, it must be the case that the conscious perspective which apprehends X at that moment does not itself seem disunified (in relation to the X being apprehended). Now what is conscious apprehending but just another description for witnessing? The witness-consciousness would hence itself, at that moment, not seem disunified in relation to the disunified X that appears to it. It follows that witnessing could never seem disunified at a time, since witnessing would have to seem unified in order to apprehend its own disunity, which is impossible.

Awareness also satisfies the criterion of being a phenomenally felt unity – a felt singularity behind any multiplicity – by virtue of its subjective sense of presence. The fact that witnessing can never appear as disunified at a time

is not, hence, to be interpreted as evidence of a mere unifying absence or noumenon (which would also never appear as disunified at a time since there would be no appearance *per se*). It is a positive presence that is felt to be unified at a time. This sense of positive presence would account for the *sense* of singularity behind the multiplicity; it would pinpoint what is puzzling about the phenomenon of synchronic unity.

From these considerations we can infer that the type of consciousness Bayne and Chalmers allude to when they speak of 'phenomenal subsumptive consciousness' is that of awareness – taking into account the fact that they would view such awareness as being part of a more complex conscious state with objects. Whether construed as part of a more complex conscious state or as a mode of consciousness that could occur on its own, awareness specifically captures the apparent singularity behind the multiplicity; the concept is fittingly implicated in the puzzle of synchronic unity. The concept of synchronic unity thus implies the concept of awareness and the concept of awareness implies the concept of synchronic unity.

2.3. Unbroken and invariable unity

When the unity of consciousness is conceived of as persisting in relation to successively changing objects, then this gives rise to a commonly held intuition. From the first-person perspective, it would seem to be the very same unchanging consciousness that unifies 'the stream' of conscious states and their changing contents. This background consciousness appears unbroken rather than gappy in its presence, somehow belying the successive flow of objects that bestow a sense of time passing. This consciousness does not itself seem to flow with time (hence the term 'diachronic' can mislead) but appears merely present to whatever objects (thoughts, perceptions, memories, sensations) pass in and out of it. The puzzle is not to do with long-term personal identity (although in a later chapter they will be related) but to do with moment-to-moment unity of consciousness: an apparently unbroken seam of present-moment presence. As before, the puzzle arises from a phenomenological tension; while no such principle of 'persisting' unity is directly observable in one's conscious life, one has a distinct phenomenal sense that such unity exists.

> Unbroken, invariable unity: Whenever objects of consciousness are apprehended as changing from one moment to the next, there appears to be an elusive yet unifying, unbroken consciousness that observes the change but does not itself change.

What concept of consciousness is implicated here?

Explaining what accounted for the intuition behind the puzzle was a problem that Hume famously grappled with. After trying to explain this

intuition by appealing to such factors as memory and imagination, he concluded in an afterword:

> But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought and consciousness. I cannot discover any theory, which gives me satisfaction on this head …
> (Hume, 1740, 175)

Hume could not abandon his intuition that, although objects of consciousness (such as perceptions) are changing, there at the same time seems to be an underlying unifying principle that persists, consciously registering the change, not itself changing. We can surmise that Hume was seeking a conscious state or mode that involved 'simplicity' as he called it – uniting the disparate elements of consciousness to a single subject of experience. We can also surmise that the state Hume sought was elusive to his introspective gaze – otherwise he would have located it among his flux of rapidly altering percepts.

I have already argued that the concept of awareness is implicated in both the puzzles of elusiveness and synchronic unity. The elusiveness dimension would accord with Hume's failure to locate unity among the discontinuous, changing objects of consciousness. The phenomenal dimension, viz., a subjective sense of unifying presence, would accord with his persisting intuition that even though he could not locate the source of unity, any account that ignores it would be incomplete. Now given that the feature of synchronic unity has to be described with reference to the concept of awareness, there is a strong initial reason to suppose that the features of unbroken invariable unity should also involve reference to awareness. But the dimensions of unbrokenness and invariability are yet to be independently established as being linked to awareness. Because these features appear to qualify the principle of unity and because 'synchronic unity' is articulated with reference to awareness, it will be most parsimonious – and in keeping with how things seem – to also specify the notions of unbrokenness and invariability with reference to awareness.

Hence the strategy here will differ from that adopted when articulating 'synchronic unity' and 'elusiveness'. Instead of first seeing whether rival concepts of consciousness can be appealed to in describing what is meant by the terms 'unbrokenness' and 'invariability', I will jump straight to the exercise of seeing whether awareness is implicated in these specified features. If awareness is implicated (and awareness implicates them), then we will have prior parsimonious reason to select the concept of awareness as depicting the kind of consciousness that seems to unite the flow of observable phenomena.

Unbrokenness: It was argued in Chapter 4 that while a subject is awake, there is an aspect to its conscious life, showing through in the self-sense,

that appears elusive, unifying and *unbroken* in its presence (endorsed by the existence of a puzzle about this feature). If 'awareness' is to capture this apparently unbroken or non-gappy aspect to conscious life, then awareness must never, in a given stretch of consciousness time, be able to present as an aspect of conscious life that is broken or gappy. If awareness is going to present itself at all, in a given conscious stretch, it has to present as unbroken in its presence. And we can quickly determine that awareness can never seem gappy in this way. As the most general mode of conscious apprehending, a subject's witnessing will be 'set up' to notice comings and goings of various mental objects, including different modes of sensing and perception – never its own coming and going. To say that witnessing directly notices its own absence, in a given state of apparently unbroken consciousness, is to commit to a contradiction. For what could possibly notice the absence of witnessing, from the first-person perspective, except a subject's witnessing? Witnessing would have to be present to notice its absence! Therefore, awareness can never be something that seems to come and go in any given state of apparently unbroken consciousness. (This is not to deny that from an 'objective' perspective, witnessing could actually be inferred to arise and pass away; but, as mentioned before, the aim is not to annex awareness to overtly objective conceptions of unbrokenness.)

The awareness also does not equate to a mere absence, which would also not present an impression of gappyness. The subjective sense of presence bestows the unbroken witnessing with a positive 'lived' dimension, which would have (non-attentively) alerted Hume to its existence. Awareness will thus seem to *be* that unbroken consciousness alluded to in the puzzle of unbroken unity.

Invariability: It was argued in Chapter 4 that there seems to be a subtle aspect to conscious life, showing through the self-sense, that is not only unbroken in its presence, but unchangeable in its quality. In a given conscious state, could awareness ever seem, essentially, to vary in its quality? On the face of it, awareness might seem to vary along with its contents. After all, within the spectrum of conscious states that may be had from one moment to the next, there can be bright attentive witnessing and dim inattentive witnessing, perceptual witnessing (through five different senses), cognitive witnessing (to do with thinking, remembering, imagining, having emotions) and introspective witnessing (to do with noticing one's own thought-processes, etc.). But to deduce from this the fact that witnessing must seem variable in its essence is to overlook the possibility of a phenomenal factor common to this set of apprehendings. We wish to know whether the concept of awareness yields such a phenomenal element that seems not only unbroken in its conscious presence, while one is awake, but unbroken in an unchanging, invariable fashion.

Our definition of awareness does indeed yield such an aspect. First, its witnessing dimension is defined as the mode-neutral aspect of *knowing* that

unites particular modes of conscious apprehension (such as seeing, hearing, thinking) into one category. Such knowing, which enables us to discern the specific *mode* of apprehension being enacted (e.g., seeing, hearing, thinking), cannot *itself* be of a quality that is particular to such a mode. Second, since awareness cannot borrow its intrinsic phenomenal character from any observable (and changeable) objects or modes of awareness, the subjective sense of presence, which seems unifying and present-centred, must appear elusively (to a person) as the *same* neutral quality in any mode of conscious apprehension. Despite changes in modes of perception and so forth, there are no discernable phenomenal parameters along which, in a given conscious state, this essential quality *could* seem to alter from one moment to the next. It is useful to compare awareness, in this capacity, to the beam of a flashlight. Although the beam may vary in intensity, either at one time or over time, the essential quality of luminosity stays the same. Luminosity is the deeper quality that unites the focused beam and its dimmer circle of surrounding light, just as the subjective sense of presence (unified and elusive) is a deeper quality that would unite each mode and moment of witnessing. And while the flashlight beam with its surrounding circle of dimmer light is a suitable analogy for attentive versus inattentive witnessing, luminous spectral colours once again may offer a suitable analogy for the different perceptual, cognitive and introspective modes of awareness. For despite their diversity, the spectral colours are also united by the deeper fact of their luminosity.

From the considerations in this and the previous section, we can thus conclude that reference to awareness, in virtue of its elusive, unifying, witnessing sense of presence, seems well suited to capture those elusively sensed dimensions of unbrokenness and invariability. Given that awareness is already implicated in the specification of elusive and (synchronically) unified consciousness – features that seem further qualified by unbrokenness and invariability – it is most economical to suppose that awareness is the single concept of consciousness that captures all these features. 'Consciousness' can thus be substituted by 'awareness' in the puzzle-generating phrase 'an elusive, yet unifying, unbroken, unchanging *consciousness* that observes the change'. The features of unbroken and invariable unity, as they purport to qualify ordinary conscious states, must be specified with reference to awareness. And awareness – that can never, from one conscious moment to the next, present as gappy or variable (unlike objects of consciousness) – must in turn be specified with reference to the features of unbroken and invariable unity.

2.4. Unconstructedness

While no particular puzzles are associated with the feature that I refer to as 'unconstructedness', this feature is implicit in each of those we have just considered. For example, just as elusiveness of the subject would not seem particularly puzzling if the subject's *modus operandi* lacked an experiential

dimension (as with a video camera), it would equally not seem puzzling if the *modus operandi* appeared as just another attendable object, for example, as a thought or perception. And that is how the *modus operandi* of subject *would* appear if it were to present as being overtly constructed: there would be the impression of its phenomenal dimension being contributed to by various thoughts or perceptions. On such a scenario, there would not, in effect, *be* an elusiveness puzzle. It is by virtue of the very fact that a subject's *modus operandi* seems unconstructed by any (attendable) objects – and yet experienced alongside them – that there is a puzzle of elusiveness.

With synchronic unity, a similar story applies. If the 'singular encompassing state of consciousness' appeared as an obvious construct – overtly contributed to by various (attendable) thoughts and perceptions – then there would be no puzzle of synchronic unity. It is only because that singular unifying state does not present as an object of consciousness – but yet seems experienced – that synchronic unity becomes a deeply puzzling phenomenon. The 'synchronic unity of consciousness', at least as formulated through the puzzle articulated in this chapter, thus involves the impression of being unconstructed by any attendable content of thought or perception. The unbrokenness and invariability of consciousness, besides seeming elusive and unifying (and hence unconstructed from this angle), are also puzzling by virtue of the fact that the consciousness they appear to qualify seems unconstructed. For it does not introspectively seem as if any overtly *constructed* content of consciousness – such as that pertaining to thoughts and perceptions – is either unbroken or invariable. Any attentive observation (such as Hume's famous introspection) will reveal such content to alter from one moment to the next. Hence the consciousness that seems unbroken and invariable will not seem constructed, but unconstructed. Given that awareness implicates the features of elusiveness, unity, unbrokenness and invariability, and that the specification of these puzzling features would imply their status as unconstructed (were they to exist), then it follows that awareness, if it existed, would also have to be unconstructed.

3. Awareness as central to phenomenal consciousness

I hope to have demonstrated that the concept of awareness is unavoidably appealed to when articulating several classic puzzles of phenomenal consciousness. These puzzles draw their quizzical nature from the seemingly unconstructed features of elusiveness, unity, unbrokenness and invariability as they appear in ordinary conscious states. In turn, the concept of awareness must be construed with reference to the aforementioned features as they present in ordinary conscious states. When appearing in a conscious state, awareness can thus never present as other than elusive, unified, unbroken, invariable and unconstructed. Given that these features figure in puzzles about the subjective or phenomenal nature of consciousness,

awareness, characterised by these features, will therefore be a necessary component of phenomenal consciousness. From this it will follow, as a corollary in the philosophy of mind, that any complete theory of phenomenal consciousness – including how consciousness (should it exist) is placed in the physical world – must address the concept of awareness.

4. Why boundedness is not implied by awareness

Having established that a conceptual link exists between the concept of awareness and those delineating the features Buddhism regards as essential to awareness, the next step is to show that the link does not extend to the concept of boundedness as played out through the role 'personal owner'. If the concept of awareness did imply the concept of boundedness in the same way that it does the concepts of elusiveness, etc., then the Buddhist account of consciousness and no-self could not be correct. On the Buddhist position, the self is deemed constructed in terms of its boundedness through the role 'personal owner', with unconstructed input from those features of consciousness that we have co-specified with awareness. Should the concept of boundedness also be implied by the concept of awareness, then the ontological status possessed by awareness and its co-specifying features – whether as unconstructed, constructed or illusory – would be conferred to boundedness. With boundedness (through the role of personal owner) on the same ontological footing as awareness *et al.*, there could be no hope of defending an account of no-self where the contributing consciousness (viz., co-specified awareness) is unconstructed, while boundedness and personal ownership are constructed. In this final section of the chapter, I set out to demonstrate that the concept of boundedness does not follow from the concept of awareness. If successful, this will leave the door open to a Buddhist account of consciousness and no-self.

I have already argued, in Chapter 4, that the impression of being a bounded, separate, ontologically unique entity is best evidenced through an assumed self-identity whose broadest mode is played out through the role of personal owner. Through identifying as a personal owner (and hence minimally with a perspectival owner), the subject assumes itself to be a personalised 'me' with reciprocal feelings of 'mine' towards various other objects that relate to the identity-role. The feeling of being a unique and personal owner qua 'me' will, should the self be real, be most suitably reflective of the self's ontologically unique, bounded status.

What has to be argued, therefore, is that the concept of awareness does not imply the concept of boundedness as it is borne out through this role of personal owner. Boundedness will of course implicate awareness. If it did not, then there could not be any sense of a bounded self. By its very definition, the self is a personalised, bounded subject that is elusive, unified and

so forth. We want to know whether the (co-specified) awareness would, in and of itself, necessitate boundedness. If we can conceive of awareness without a sense of boundedness, then this will be enough to demonstrate that the concept of awareness does not imply the concept of boundedness.

It can be easily demonstrated that the concept of awareness does not imply that of boundedness. While not yet having investigated the psychological possibility of awareness *sans* a sense of personal boundedness, we can certainly conceive of this possibility, which is all that is needed for these purposes. In Chapters 3 and 4, I considered various pathologies (anosognosia and depersonalisation) involving subjects to which a portion of perspectivally owned objects did not, at a time, seem personally owned. All we have to conceive of is a 'global' case where, at a time, no perspectivally (or possessively) owned objects seem personally owned by its subject – and from the analysis of Chapter 4, this seems easy enough to do. From the analysis, we can infer that with no sense of personal ownership towards any items, there would be no reciprocal sense of being a personal owner, and with no sense of (personal) 'me' and 'mine', there would be no sense of boundedness. The remaining perspectival ownership would nevertheless imply the presence of awareness. Newborn infants and primitive organisms (as well as *Arahants*) can conceivably exemplify such cases of awareness *sans* a sense of boundedness. Now I have not yet discussed whether, as a matter of psychological fact, such cases exist – or whether a loss in the sense of boundedness would entail a loss of actual boundedness. But the very fact that we can conceive of and indeed investigate the psychological possibility of awareness without a sense of boundedness, shows that the concept of awareness does not imply the concept of boundedness as it does the concepts of elusiveness, unity and so forth.

Conclusion

In this chapter I have sought to demonstrate, in accordance with Buddhism, that the concept of awareness, viz., witnessing with a subjective sense of presence, is to be co-specified through the concepts of elusiveness, unity, unbrokenness, invariability and unconstructedness. I have also sought to demonstrate, in accordance with Buddhism, that the concept of awareness does not similarly imply the concept of boundedness that is associated with the self's central role of personal owner. But I have not as yet demonstrated, in accordance with Buddhism, the independent reality of awareness or its specifying features. For this to be established, awareness must be shown to exist in the manner that it purports to exist – as non-illusory and unconstructed. The next chapter will thus present an argument that the co-specified awareness, integral to our sense of self, is non-illusory – and that this non-illusory status implicates its unconstructed status.

7
The Unconstructed Reality of Awareness

Introduction

In this chapter, I attempt to establish the non-illusory and independently real status of awareness and its intrinsic features (whose concepts are implied by the concept of awareness). For this non-illusory status to be established (which is needed if awareness is to have the independent reality ascribed to it by Buddhism), awareness must be shown to exist in the manner it purports to exist. Awareness purports to exist as a witnessing presence that is unified, unbroken and yet elusive to direct observation. As something whose phenomenology purports to be unborrowed from objects of consciousness, awareness, if it exists, must exist as *completely unconstructed* by the content of any perspectivally ownable objects such as thoughts, emotions or perceptions. If *apparent* awareness, perhaps by virtue of one or more of its defining features (that form part of its content or 'aboutness'), turned out to owe its existence to such object-content rather than to (unconstructed) *awareness itself*, then that would render awareness constructed and illusory and hence lacking in independent reality. For example, if the analyses of 'no-self' by Western thinkers mentioned in Chapter 5 (such as Hume (1739) and James (1890)) were correct, with the features of unity, unbrokenness and invariability being constructed, then awareness, co-defined with these features, would similarly have to be viewed as being constructed and hence illusory. In this chapter, I first argue that the co-specified awareness, central to our sense of self, has unconstructed status. I then address a possible objection from those who may still deny the independent reality of awareness by advocating a version of eliminative materialism. Finally, I consider the influence of what I call the 'object-knowledge thesis', which has, in my opinion, thwarted the popularity of awareness as a concept in the philosophy of mind.

1. The central argument

It has been established that if awareness is a construct, then it must be an illusion. If it can be shown that awareness is not (and perhaps cannot

possibly be) an illusion, then it will follow that awareness is not (or cannot possibly be) a construct. To see that awareness is not an illusion we need to first review what, ontologically speaking, is involved in the generation of an illusion. In Chapter 5, I distinguished between two common usages of the term 'illusion', namely, (a) illusion *qua* vehicle, viz., appearance-of-content-to-person-P and (b) illusion *qua* content-of-appearance-to-P. Now from the outset I defined 'illusory X-as-F' as an X that, from the perspective of P, purports to exist in a manner F when it does not actually exist in manner F. It should hence be clear that 'illusory X-as-F' pertains to (b), namely, to illusion-qua-content rather than to (a), namely, illusion-qua-vehicle. Indeed, unless we accepted the non-illusory status of the illusion-vehicle – the *appearance*-of-(X-as-F)-to-P – there could be no such phenomenon as *content*-of-appearance, X-as-F, which turns out to fall short of its purported mode of existence. In other words, there has to be a non-illusory appearance as an event occurring to person P at time t, in order for there to *be* any content of that appearance conveyed to P which could, on further examination, turn out to be illusory. For example, in the case of the Müller–Lyer Illusion, the visual presentation of two objectively uneven lines to a subject is not illusory; what is illusory is the conveyed content <two objectively uneven lines>. And with the self, what various thinkers deem illusory is not the feeling of being a self (viz., the appearance or sense of self), but the self (the content), which that feeling conveys. The *appearance*-of-(X-as-F)-to-P-at-t is hence not illusory in our defined sense; it exists as the vehicle *for* the illusory content.

Given this, what are we to say about awareness, with the features in terms of which it is specified? Like the self, awareness is peculiar in that its appearance-vehicle and content seem, from the perspective of a person, to coincide; there does not, phenomenologically speaking, seem to be a *gap* between the content that awareness conveys (viz., a witnessing unified, elusive, unbroken, invariable sense of presence with unconstructed status) and the vehicle of subjective appearance which conveys that content. From this, we may be tempted to conclude that since the *appearance* of awareness must be non-illusory, then for that very reason the content <awareness> (with all its co-specifying features), which seems to transparently characterise the nature of that appearance itself, must likewise be non-illusory. This move, however, is fallacious. Appearances can be deceptive. While certain content (such as unconstructedness) may phenomenologically *seem* to intrinsically characterise an appearance, the appearance may conceivably turn out, intrinsically, to lack that feature (which would render the feature/content illusory). After all, if we were to conclude that awareness is non-illusory (and hence unconstructed, as it seems to be) on the simple basis of how it phenomenologically appears to the average person (as united with its content), then we should also have to conclude that *the self*, whose very presentation seems to suffuse awareness with the content <boundedness>, is non-illusory. In the following chapter,

I hope to demonstrate how boundedness does not intrinsically characterise the ontology or appearance of awareness – even though it commonly *seems*, from the first-person perspective, to characterise awareness in this fashion. (I have already shown that the concept of boundedness is not in fact implied by the concept of awareness.)

Nevertheless we may still ask whether there is some *particular* feature to awareness whose very existence is implicated in the ontology of an appearance, such that without this feature there could not *be* an appearance – including the appearance-vehicle for an illusion. When we put the question this way, it becomes apparent that there is at least one such feature to awareness, namely, the central witnessing component implicated in the 'appearing-to-person-P-at-t' part of the illusion-vehicle. How do we get from 'person' to 'witnessing'? Simply by noting that in order for any illusion to take hold, there must be a person to whom the illusion appears. The person must be conscious and cognisant of the content, such that he can be fooled by it. He must thus have a *perspective* through which the content is consciously *apprehended*, regardless of the mode of apprehension. He is always able to tell, moreover, the mode of apprehension (e.g., visual, auditory or cognitive) in which an illusion is being presented. But this mode-neutral conscious apprehension from a perspective adds up to none other than a *subject*, namely, to *witnessing-from-a-perspective*.

Witnessing is thus built into the very appearance-vehicle that would *carry* any illusory content, and as such, it cannot be equated *with* illusory content. Witnessing is a necessary condition for the very possibility of an illusion-vehicle. Now we gathered from Chapter 6 that witnessing, from a first-person perspective, cannot present as other than elusive, unified, unbroken, invariable and unconstructed – at least to the extent that these features seem apparent in ordinary conscious states. But the delimiting extent to which they *seem apparent* implies a dimension of *appearance* to the nature of witnessing and its features – the appearance being, of course, to do with the subjective sense of presence. The features of elusiveness and so forth are not thus merely inferred to exist through abstract reasoning (à la Kant); they seem immediately *alive* to us in ordinary conscious states, through the subjective sense of presence. They seem to *qualify* the subjective sense of presence such that it is *witnessing presence* – not merely the witnessing by itself – that appears inextricably elusive, unified, unbroken and so forth. Witnessing and its features are therefore implicitly presented *through* the subjective sense of presence – the elusive presence at the root of those phenomenological tensions behind the puzzles (the puzzle being that the features seem present, yet they cannot be attended to). Given that witnessing is necessary for the possibility of an illusion-vehicle, and that its qualifying features imply (through their felt limitations) a subjective sense of presence, it follows that *awareness* – witnessing presence conceptually characterised as elusive, unified, unbroken, invariable and unconstructed – is

also a necessary condition for the possibility of an illusion-vehicle. Awareness cannot, therefore, be the content of an illusion and hence, it cannot be a mental construct (for we determined that if awareness were constructed it would have to be illusory). All the features that awareness intrinsically brings to the sense of self will thus be non-illusory. This implication notably does *not* extend to boundedness which, as we argued in Chapter 6, is not implied by the concept of awareness (although it would usually *seem*, phenomenologically, to be intrinsic to the awareness we experience).

Awareness therefore exists, intrinsically speaking, in the manner that it purports to exist, namely, as the content of an appearance in which it is not contributed to by any perspectivally ownable objects such as thoughts, feelings, emotions or perceptions. The unconstructed awareness really does underlie its perspectivally ownable objects (as it seems to do); the objects do not construct awareness. The claims made by Hume (1739), James (1890), Dennett (1991), Flanagan (1992) and Damasio (1999), to the effect that unity, unbrokenness and invariability are constructed by such objects, are therefore false.

2. The spectre of eliminative materialism

I have argued that awareness can be neither illusory nor constructed. Some philosophers may nevertheless claim that *no* subjective or mental phenomena can have any reality *because* awareness, implicated in all subjective phenomena, is not real. It is hard to see how such a view could be developed unless it was, in the end, some version of eliminative materialism. Defenders of this position may claim that subjective witnessing, along with all mental phenomena (including illusion-vehicles), is a byproduct of the brain, lacking in any reality. However, there are serious problems with this position, a position that also happens nowadays to be rather unpopular in the philosophy of mind. It is unclear, without resorting to dogma, how one could construe the manner in which awareness and its subjective (perspectivally owned) contents could lack reality. We have demonstrated that it is not an option for anyone, including the eliminativist, to declare awareness to be either illusory or constructed – two major ways in which something can, by parties of differing metaphysical commitments, be plausibly deemed to lack reality. This means that if awareness were to lack reality, it could not do so in such a way that its actual ontological status would conflict with its purported mode of existence.

The eliminativist might nevertheless insist that awareness *is* illusory since its very appearance, as purporting to exist *simpliciter*, conflicts with the eliminativist standard of reality, which denies existence to appearances. This response, however, would fail to peg awareness as an *illusion* – even if the eliminativist standard of reality turned out to be correct. It would first of all unacceptably alter the notion of 'illusion' such that the structure of an

illusion would no longer be confined to that of a non-illusory appearance-vehicle that carries illusory content. The vehicle *of* the illusion would be illusory! Even if we were to accept this distortion to the term 'illusion', awareness, if illusory on this count, would have to *purport to exist* with reference to some standard of reality relevant to that of eliminative materialism. The standard of reality relevant to the eliminative materialist is usually one whose acceptable ontology, for various theoretical reasons, excludes mental phenomena such as appearances. But the phenomenon of awareness is silent with reference to its wider metaphysical underpinnings – it does not wear on its sleeve 'purports to have independent (rather than just apparent) reality in a system of metaphysics that excludes the existence of mental phenomena'. The unconstructed manner in which awareness purports to exist is simply as the intrinsic content of an appearance (viz., that of unified, elusive witnessing, etc.) by which it is not contributed to by perspectively ownable objects such as thoughts and emotions. Awareness cannot, therefore, be deemed illusory by purporting to exist in a manner that relates to the agenda of the eliminativist materialist.

If awareness cannot be an illusion, then a possible step towards denying reality to awareness may be to argue that awareness (along with all other mental phenomena) is epiphenomenal, lacking in any genuine causal power (all the causal work being done by underlying neurology). If awareness were epiphenomenal, then this would *not* render it illusory, since awareness (a) cannot be illusory for reasons already mentioned, and more specifically, (b) does not 'wear on its sleeve' any story about its metaphysical underpinning, including the ultimate status of its causal efficacy in relation to the brain. The concept of awareness would suffer no internal contradiction from the possible scenario of awareness supervening upon neurological properties which do any causally relevant work. Suppose that awareness were epiphenomenal in this way. This would not be enough to guarantee the non-reality of awareness, since awareness could still plausibly be construed to exist as a real-enough *effect* of the neurological states.[1] There would need to be an additional claim, which declares any causally impotent X to lack reality. One may for instance stipulate, as Richard Baron (2000) does, that only causally active, objectively measurable elements have a place in the elite 'scientific worldview' and that anything which is not a part of this scientific framework lacks reality. The problem with this approach is that it seems to violate the sensible philosophical dictum: *preserve appearances unless there is overriding reason not to*. To my knowledge, eliminative materialism does not provide overriding reason to renounce what seems so obvious, namely, that subjective phenomena, epiphenomenal or not, are as much a part of the world as 'scientifically respectable' properties. It is probably for this reason that eliminative materialism never gained popularity in the philosophy of mind. Philosophers of mind nowadays are generally more interested in explaining than denying the reality of mental

phenomena. The history of philosophy has shown that repeated attempts to banish subjectivity to the 'attic of the mind' because it is methodologically difficult to deal with, has not made subjectivity – including the witnessing perspective from which we approach the world and its scientifically respectable properties – disappear. The mind–body problem is still at large, and facing up to the reality of awareness – whether epiphenomenal or not – seems a necessary step in resolving the mind–body problem.

3. The object-knowledge thesis

While there have been some who have advocated the importance of the concept of awareness (or something similar), Western philosophy of mind has not been that receptive to the concept of awareness *as defined in this book*. The reluctance to give a central place to some version of the concept in the philosophy of mind has partly been because awareness is elusive and difficult to define and – perhaps the other side of the coin – at odds with a tacit but widespread assumption that I shall call the *object-knowledge thesis*. The object-knowledge thesis states that all knowledge and experience must be derived from the object side of the apparent subject/object dichotomy (remembering that objects are, in principle, able to be attended to). It rules out the possibility of unmediated knowledge or experience that pertains specifically to the subject side of the apparent subject–object dichotomy – in particular, to its *modus operandi* of awareness. With an epistemic tradition based on a bifurcation of rationalism and empiricism, it is unsurprising that the object-knowledge thesis often goes unchallenged in Western philosophy. Despite their differences, both empiricism and rationalism rest upon the object-knowledge thesis. Empiricism maintains that all things we can know must ultimately derive from deliverances of the five senses; rationalism holds that not all things we can know must derive from the five senses but can be contributed to by the intellect. Both empiricism and rationalism limit the source-pool of knowledge and experience to deliverances from either the senses or the intellect, hence ruling out the possibility of unmediated, experiential contribution from the subject's *modus operandi*. (It should be evident that the concept of 'knowledge' is to be understood *synecdochically*, that is, as representing a broad group of related epistemic notions such as apprehending, experiencing, intuiting and so forth.)

I think it is not an exaggeration to say that part of what has been so puzzling about problems of consciousness (relating to its subjectivity) has been an unquestioning allegiance to the object-knowledge thesis. In other words, the adoption of this thesis has made the problems harder than they have to be. For instance, Hume (1739), locked into his empiricism, could not countenance the possibility that his intuitions of simplicity (unity) and identity (uninterruptedness and invariability) could be grounded in anything other than the products and mechanism of his empirical perception.

Had he admitted to the reality of awareness as their source, he could have moved forward in his explanation of these intuitions, rather than admitting defeat. And Kant (1787), while allowing that unity did have a source in the noumenal (non-empirical) subject, did not bestow the subject with any intrinsic phenomenal character that could be brought into conscious experience. Kant's analysis did not therefore account for the persisting intuition that there is *something it is like* to be in a *unified* state of consciousness whose phenomenal character is not entirely reducible to features of the objects unified. Unity is not, à la Kant, merely knowable to the minds of philosophers as the conclusion of some transcendental inference that must hold if conscious experience is to be possible. Unity is elusively apparent *to* first-person experience – everyone's experience – and it is *because* it is apparent that it has continued to puzzle philosophers, especially those who adhere to the object-knowledge thesis.

There are nevertheless some concepts of consciousness besides awareness that may imply a rejection of the object-knowledge thesis. In Chapter 6, I made mention of concepts of consciousness that, following Aristotle and Brentano, have been recently developed by thinkers such as Victor Caston (2002), Uriah Kriegel (2004) and Greg Janzen (forthcoming). The similarity that these concepts bear to that of awareness is that the phenomenal character of the referred-to consciousness, on these theories, is taken seriously and is recognised as not being exhausted by the character of those (primary) objects to which the consciousness is directed at a given moment. The phenomenal character of consciousness is also said to be accounted for by consciousness inattentively taking itself as its own reflexive 'on the side' *object*. And here is where divergence can occur. Following on from the comments in Chapter 6, we can infer that if 'object' is to be taken as implying *attendability* (as I use the term), then the object-knowledge thesis is not after all rejected by advocates of the Aristotelian/Brentanian theory. For it will be held that consciousness gets its phenomenal feel from taking itself as something that *could* be attended to (even if most of the time it is not – see Kriegel (2004)).[2] But if 'object' is taken to convey the broader notion of *content* – as in consciousness being *about* itself – then the Brentanian/Aristotelian concept of consciousness may well, like the concept of awareness, deny the object-knowledge thesis; it may in fact *incorporate* the notion of awareness (keeping in mind its additional object-directedness), for content need not be attendable – as my Chapter 5 definitions reflect. Awareness comes out as being *unconstructed*, on my definition, because it is the *content* of an appearance by which it is *not* contributed to by such objects as thoughts and perceptions.

A recent philosophical position that could also serve to challenge the object-knowledge thesis is that developed through the notion of 'phenomenal subsumptive consciousness', proposed by Bayne and Chalmers (2003) and discussed in Chapter 6. Their position seems at odds with the object-knowledge

thesis. For, when unpacked, it states that there is something it is like, at a time, to be in a single unified state of consciousness, and that this 'something it is like' is not reducible to merely the phenomenal characters of the objects subsumed by it. Phenomenal subsumptive consciousness seems to add its own subjective character to conscious states hence implicating, as I said in Chapter 6, the notion of awareness.

I would suggest, furthermore, that the concept of awareness deserves a central place in the philosophy of mind, not only because it captures an aspect of consciousness common to these puzzles about subjectivity. Accepting the concept of awareness as centr al also goes some way towards *solving* the puzzles. We have seen that a persistent obstacle to the solution of the 'elusiveness' and 'unity' puzzles has been the assumed object-knowledge thesis, which does not take seriously the possibility of unmediated subject-knowledge, and hence awareness. Accepting the reality of awareness involves a rejection of this common assumption and hence the removal of a common obstacle, permitting one to freely acknowledge what is behind these puzzles in the first place. It also has a significant bearing on what is known as the 'hard problem of consciousness', whose solution aims to explain the place of consciousness in the world. Tackling the hard problem must first involve defining what consciousness *is* and I hope to have argued convincingly that consciousness involves the quintessentially subjective dimension that is awareness. (The subjectivity of consciousness is what allegedly makes it a *hard* problem.) I believe that once the object-knowledge thesis is rejected, and the reality of awareness accepted, better progress will be made in dealing with the hard problem of consciousness. For any explanation of consciousness *qua* awareness will simultaneously impact upon the phenomena of elusiveness, unity, unbrokenness and invariability, thus accounting for several puzzling features at once.

8
How the Self Could Be a Construct

Introduction

This chapter tackles those questions at the heart of this book: does the self we assume we are have independent reality? Or is it a construct and illusion? If it is a construct and illusion, then how is it put together?

I have argued that a number of features we implicitly *ascribe* to our selves, through our modes of thinking and living, are not constructed or illusory. The features are awareness and its qualifying attributes, all of which purport to be unconstructed: unity, elusiveness, unbrokenness and invariability.

So if the self lacks reality, it is probably due to its boundedness. Boundedness is that aspect by which a subject's awareness presents itself as part and parcel of an *ontologically unique* subject with personalised boundaries that separate a *me*, viz., a self, from the rest of the world. It was argued in Chapter 4 that our *sense* of being a bounded entity, viz., the subject's sense of being *awareness-as-bounded*, is most closely tied up with the role 'personal owner' – a role that must be assumed before one can identify as a thinker of thoughts or agent of actions.

It has already been demonstrated that the concept of boundedness does not, unlike the concepts of unity, elusiveness and so forth, follow from the concept of awareness, since it is possible to conceive of awareness without boundedness. This suffices to show that, on a *prima facie* level, awareness is not as tightly associated with boundedness as it is with elusiveness, unity and so forth. The fact that we can conceive of awareness without boundedness, however, does not show that awareness is not, as a matter of psychological fact, intrinsically bounded through the role of personal owner. Conceivability does not imply actuality. Perhaps, awareness must in fact occur as part and parcel of an unconstructed, bounded, personal owner – so as a self – a fact not obvious through merely analysing the concept of awareness. Perhaps, in other words, the bounded self (equivalent to awareness-as-bounded) is unconstructed as it purports to be, and hence not illusory.

In this chapter, I investigate evidence to the contrary, which suggests, I think convincingly, that the bounded-self-as-unconstructed, by virtue of its boundedness, is in fact *not* unconstructed – as it purports to be – but constructed. Such evidence will be drawn from findings and ideas of neurologist Antonio Damasio. In view of the analysis in Chapter 5, I will thus be arguing that the self, which purports to be the content of an appearance by which it is *not* contributed to by perspectivally ownable objects, *is in reality,* the content of an appearance by which *it is* contributed to by such objects. Such objects will comprise mainly of thoughts and emotions that are driven by a boundary-promoting concern for the welfare of one's own person, related to that aspect of 'the self' which seeks happiness and avoids suffering. (Although Buddhism would regard the happiness-urge to have some non-illusory roots in *nibbāna* – something not argued for in this book – they would also take it to involve the mistaken belief that such an urge can be fully satisfied in the conditioned world, which ties it to *taṇhā* and boundedness.)

Following this, I will propose a positive theory of how the illusory self could arise from the input of two tiers: awareness (qualified with the various features) and those boundary-promoting (*taṇhā*-driven) thoughts and emotions. The theory will recruit the process of identification that was described in Chapter 4.

1. Revisiting evidence that awareness purports to be a bounded self

Boundedness is that feature by which a subject's awareness comes 'dressed up to the party of life' as a specific identity with boundaries – as this *very* self that is ontologically unique, separate from the rest of the world. Each co-defining feature of the witnessing presence seems to partake of this 'dressing-up', such that the features seem to qualify not merely the bare, impersonal witnessing presence, but the bounded personal self as a whole. In light of Chapters 6 and 7, which provided some detailed discussion of the features of awareness, it is worth at this stage being reminded of how it *seems* that a *bounded self-entity* displays those very features (thus reiterating the theme of Chapter 4).

It is not, hence, mere bare awareness, but a *personalised bounded self* that seems to view the world from a unique first-person perspective. The perspective of the witnessing presence thus appears to be the *self's* perspective. Moreover the personalised self, not just bare awareness, seems to *be* that entity to which, at a given conscious moment, one's thoughts and perceptions seem unified. Even when thinking about aspects of one's personality, it is always, as David Velleman (2002, 114) puts it, *with* those parts of the personality being currently identified with as 'me'. Those very parts of the personality with which I am currently thinking will never appear as separate, observable objects of consciousness; they always appear to be *at one*

with the witnessing presence who is observing or thinking. The observing *self* would thus appear to also inherit the observational elusiveness of witnessing presence as well as its unity. It hence seems as if a personal, bounded, unified, elusive witnessing *self* – not a mere impersonal, unified, elusive, witnessing presence – is confronting the world and its objects as 'other'.

Nor does this elusive, unified witnessing self seem, from the viewpoint of the subject, to pop in and out of existence. From moment to moment, it reflexively seems to the subject as if it is the *self*, not merely impersonal awareness, that is the unbroken witness to a stream of thoughts and perceptions. It also seems to the subject as if it is a *self* that partakes in longer-term numerical identity such as when waking up from deep sleep, or when remembering an earlier stage of one's life. And despite obvious qualitative changes, it still appears to the subject as if a qualitatively invariable and yet elusive thread of 'me-ness' ties up the self's identity not only from one moment to the next, but over a lifetime. Finally, nothing about the bounded self, as reflected upon from the first-person perspective, would seem to be overtly constructed by thoughts, emotions or perceptions that are harboured at a given time. This apparent unconstructedness of the self is a direct upshot of its apparent elusiveness and unity. Thoughts *qua* thoughts (etc.) will always appear to be owned or initiated *by* the independently existing unified self, seeming to exist as objects that are observed by, and separate from, the self.

2. Is the bounded self a construct?

2.1. Can awareness exist without sense of bounded self?

Suppose that awareness must, as a matter of psychological necessity, occur with a *sense* of bounded self. The previous section suggests this to be an assumption that most of us tacitly take on board. If correct, this assumption would provide a powerful reason to suppose that awareness must, as a matter of psychological necessity, occur with – indeed *in the format of* – an actual self. If presented only ever in the bounded 'format' of a self, then it is very likely that awareness would confer its unconstructed status to the self as a whole. In this section I aim to rule out the possibility that awareness must always occur with a sense of self, by seeking an actual example, from neuropsychological literature, of awareness *sans* sense of self. We will see why typical cases of depersonalisation, sometimes offered as examples of where the sense of self is missing, do not seem to serve as good examples – but why epileptic automatism, in contrast, *is* a good example. While these findings will not by themselves establish that awareness actually does lack the boundedness of a self (which would imply the self to be a construct), they will be an important first step. If awareness need not occur with a *sense* of self, there is initial reason to suppose that it need not occur with a self, either.

The examples of two different psychopathologies, anosognosia and depersonalisation, were used in Chapter 4 to illustrate the difference between perspectival and personal ownership. Not being a trained psychologist, I am, of course, no expert on these pathologies. What I do suggest is that many available descriptions from recognised psychology manuals (available online), together with descriptions from sufferers themselves,[1] provide strong *prima facie* reason for supposing that depersonalisation (at the very least) does not involve a complete loss of the self-sense. While anosognosia involves the subject's ignorance of a specific disorder such as paralysis (such that one may deny feeling personal ownership towards one's paralysed arm), depersonalisation seems to affect the subject in a more global way. According to the DSM-IV Manual (2004, online)[2] depersonalisation involves the 'persistent or recurrent experiences of feeling detached from, and as if one is an outside observer of, one's mental processes or body'. Another psychiatric manual (2004, online) adds: 'This disorder is characterised by feelings of unreality, that your body does not belong to you, or that you are constantly in a dreamlike state'.[3] From this, it may be tempting to conclude that during typical episodes of depersonalisation, the sense of self is entirely absent, since it would appear from the definitions that feelings of personal ownership and self-identification towards one's body or mental processes will be suspended or absent. However, while not denying that the sense of self may be somewhat compromised, there is reason to question whether a sense of self will, in fact, be missing during such episodes. For the sense of self to be present at a given time, at least along the lines developed in this project, it is enough that there be *some* item – internal or external – towards which the subject harbours the reciprocal senses of 'me' (through identification as its personal owner) and 'mine' (through a sense of personal ownership towards the item). So long as there is *any* sense of personal 'me' or 'mine' with regard to *any* item, then the subject will harbour a sense of self *per se*, *even if* the subject should lack a sense of self-identification and/or personal ownership towards all other items that are, at that time, within their conscious purview.

My suspicion, with regards to depersonalisation, is therefore as follows. The cases that I have read about strongly suggest that those who suffer an episode of depersonalisation, at a given time, feel a lack of personal ownership (my-ness) towards *some* and perhaps even *most* of their perspectivally owned experiences. But there remains evidence of at least *some* such items towards which they, as subjects, harbour a sense of self-identification and/or personal ownership – which will guarantee, on the account developed in this project, that they also harbour a sense of self. The items in question invariably involve a pod of recurring negative emotions – even if blunted – that pertain to their realisation that they are in a depersonalised state: they realise there is something wrong and they *wish the state and its attendant sensations would go away*. This 'wishing the state would go away' is a clear

instance of emotional investment or what Buddhism would refer to as *taṇhā*. The negative emotions arise because the person is in a situation that he wishes was otherwise. In Chapter 4, the presence of negatively charged emotion, in the section on 'Consistent Self-Concern', was enlisted among major evidence for a subject's identification as a bounded self. *Taṇhā* was also noted as being strongly associated with all the major modes of assumed self-identity, with an increase in *taṇhā* appearing to correlate with an increased degree of identification or felt personal ownership towards various items. While a causal link between these factors has yet to be established – something I hope to do later in the chapter – this correlation should in itself serve as evidence that a sense of self is present whenever there are negative emotions or *taṇhā*. Given these reflections, it is thus an easy step, when considering such passages as the following description from the Depersonalization Research Unit, to infer a sense of self in such patients:

> Many sufferers describe it as 'terrifying', 'like losing your sense of being alive', 'a living death', 'like being detached from your own body, your loved ones, your feelings ?' People say that it is as if their mind is full of cotton wool; they pray that they will wake up and it will all be clear once more. Many describe de-realisation: as if the world around them is like a movie or that they are separated from other people by an invisible pane of glass. When such unpleasant feelings persist without explanation, the person may be judged to be suffering from depersonalisation disorder. It can be brought on by severe stress or emotional turmoil but may also appear out of the blue, and apparently suddenly.
>
> (Depersonalization Research Unit, 2006)

From this passage, I would surmise that typical sufferers of depersonalisation do not fully identify with the normal bodily sensations that are usually assumed to be integrated with the subject. Yet it would seem that they still identify as the personal owner of the unpleasant feelings: both the primary 'cotton wool' blunted feelings and the distressed emotional reaction *to* the blunted feelings (terror, etc.). Indeed if they did *not* believe the blunted feelings to be personally owned by them in the first place, then it is unlikely they would be feeling so distressed, praying that *their* condition would clear up. The high degree of emotional distress suggests, therefore, a sense of personal my-ness towards both sets of unpleasant feeling which, in turn, indicates a sense of identity as the personal owner of the feelings. I thus claim, in keeping with Chapter 4, that whenever such emotions accompany an episode of depersonalisation, the depersonalised subject is identifying itself *with* the perspectival owner of the emotions and hence *as* the personal owner of those negative emotions and hence, as a bounded self in this capacity.

Sufferers of depersonalisation also typically appear to identify with – and *value* – that part of their perspective that has the insight into what is

happening; they do not seem to be delusional about their condition. According to Los Angeles psychiatrist Oscar Janigar, 'reality testing remaining intact' is partially constitutive of the disorder's DSM-IV definition (2004).[4] The sense of terrible isolation often reported with depersonalisation is further testimony to the fact that a sense of bounded self is still there – in fact with boundaries tightened to such an abnormal degree that the subject feels as if there is an impenetrable barrier between their constricted self and the world. As one patient puts it: 'It's like I fall deep within myself. I look at my mind from within and feel both trapped and puzzled about the strangeness of my existence.' (*op cit*).

There are more extreme forms of depersonalisation disorder – the most extreme being Cotard's syndrome.[5] This complex condition cannot be summarised in a paragraph, but the primary symptom is the delusion that one is dead, implying an absence of 'intact reality testing'. While Tim Bayne (2004, 232–233) holds that cases of Cotard's syndrome are evidence that a sense of personal ownership or 'my-ness' towards one's thoughts is altogether missing, I am sceptical that this is the case. What makes me suspect that there is *not* a lack of identification as the personal owner of some X, is once again, the seeming frequency with which such cases occur with negative emotions or delusions such as those of immortality. According to the autism homepage website:

> Berrios & Luque (March 1995) performed a statistical analysis on 100 cases of Cotard's Syndrome. Cotard Syndrome was found to affect men and women in equal numbers and severity. The elderly were more likely to develop the syndrome. Depression was reported in 89% of subjects, anxiety in 65% of subjects, and guilt in 63% of the subjects. Delusions were categorized thusly: nihilistic delusions concerning the body (86%), nonexistence (69%), hypochondriacal delusions (58%), and delusions of immortality (55%).
>
> (Heffner, 2004)

If this quotation is anything to go by, I would suggest that the typical Cotard's patient does not cease to identify with things, but has *highly abnormal* patterns of identification. That is, he does not identify with a living body, as most of us do, but identifies with a dead body, appropriating the idea 'dead body' to his current witnessing perspective. Thus he believes that *he* (qua *self*) is literally dead, sometimes to the point where he can 'feel' the worms of decomposition. The negative emotions are simply in keeping with the assumed identity; believing that one is dead and rotting is not likely to be a pleasant experience (and usually prolonged negative emotions precede Cotard's syndrome). The high co-occurrence of negative emotions is, I suggest, *evidence* that the patient identifies as their personal owner, showing little doubt that he believes it is *he* who is dead. On the basis of this, I would

surmise that, as with the less extreme forms of depersonalisation, an intact (although unusual) sense of self is typically present in those who suffer from Cotard's syndrome.

There are, however, pathologies that *do* seem to provide far clearer evidence that the sense of bounded self (and hence personal ownership) can be properly lost, if just temporarily. During such episodes, there is no evidence of 'intact reality testing' and no evidence that the subject is emotionally upset about what has befallen *him*. One classic example of such an episode is known as an *epileptic automatism*, which can be brought on by a brain seizure. Antonio Damasio (1999) has observed and described in some detail the behaviour of patients who have undergone such episodes. The patient may be having a perfectly normal conversation and then suddenly without warning,

> freeze whatever other movement he was performing, and stare blankly, his eyes focused on nothing, his face devoid of any expression – a meaningless mask. The patient would remain awake …This state of suspended animation might last for as little as three seconds. … The longer it lasts, the more likely it … will be followed by absence automatism, which, once again, can take a few seconds or many. … As the patient unfreezes he looks about … his face remains a blank …he drinks from the glass on the table, smacks his lips, fumbles with his clothes, gets up, turns around, moves towards the door, opens it … then walks down the hallway. … the automatism episode would come to an end and the patient would look bewildered, wherever he would be at that moment. … would have no recollection whatsoever of the intervening time [during the episode].
>
> (1999, 96–97)

According to Damasio, the patient during the episode,

> would have remained awake and attentive enough to process the object that came next into his perceptual purview, but inasmuch as we can deduce from the situation, that is all that would go on in the mind. There would have been no plan, no forethought, no sense of an individual organism wishing, wanting, considering, believing. There would have been no sense of self, no identifiable person with a past and an anticipated future – specifically, no core self and no autobiographical self.
>
> (1999, 98)

We can note that Damasio's notion of 'core self' corresponds closely with what I have been calling the 'sense of self', viz., the basic sense of being an elusive, bounded, individual owner, thinker and actor. The ownership that concerns Damasio is clearly personal, involving a distinct sense of 'me' and 'mine'.[6] The 'autobiographical self' is simply an extension of the core self, involving a sense of one's long-term extended personal history with the

ability to imagine one's existence into the anticipated future. We may recall from Chapter 4 that another of Damasio's patients, 'David', with his 45-second memory, had an impaired autobiographical self but intact core self (Damasio, 1999, 117–121). The patient undergoing an episode of epileptic automatism, however, has not even this much window for reflection – he merely acts 'within the microcontext of the moment' (Damasio, 1999, 98). Nothing would seem to be reflected upon or committed to memory – or at least to memory that he can normally access. There is hence no evidence that he identifies with any aspect of his experience as 'me', and so no evidence that he harbours a sense of personal ownership or 'my-ness' towards his experience. In addition to this (and perhaps crucially), all signs of emotion are entirely absent – and Damasio holds, as I do, that emotions (at least those associated with 'self'-concern) are reliable indicators for consciousness with a sense of self (Damasio, 1999, 99–100). I think it is reasonable to infer, as Damasio does, that during such an episode, the *sense of a bounded individual self is absent*. I think that it is also reasonable to infer, from the fact that the patient is awake and minimally responsive to his environment, that *awareness is present*.

There are other pathologies, mentioned by Damasio, that also point to an absence of what he terms a 'core self' (equivalent to a sense of self) but with the presence of awareness: akinetic mutism and advanced cases of Alzheimer's disease (Damasio, 1999, 101–106). To describe these in detail will not add much to the discussion, but to mention them in passing is useful, since it adds weight to the empirical evidence that a person's awareness can psychologically exist without the reflexive sense of a bounded self. Awareness *sans* sense of self is not, it would seem, a mere conceptual possibility. Such cases make it more plausible to suppose that newborns and an array of primitive organisms will also harbour awareness without a sense of bounded self.

Damasio's findings would appear to indicate, then, that a person can harbour awareness without reflexively sensing this awareness to be part of a separate identity, viz., a self. This does not yet imply that awareness is not *in fact* part and parcel of a bounded self, with its feature of boundedness (enabled through identification) simply obscured or disabled on occasion by cognitive malfunction. However, this discussion *does* seem to rule out what *would* be strong evidence for an unconstructed self, namely, that of awareness always transparently exhibiting a sense of boundedness (as it does elusiveness, unity and so forth).

2.2. The next steps to arguing that the self is a construct

We have ruled out the scenario that whenever there is awareness there must be a sense of self, a scenario that would powerfully suggest the bounded self to inherit the unconstructed status of awareness. The next step in arguing for the constructed status of self will involve clarifying just how, on a psychological level, the self *could* be constructed. How *could* the content of

thoughts, emotions and so forth help create a sense of bounded self in lieu of an actual self? In the following two sections I advance the case for the self's constructed status by (a) describing a Buddhist-derived account of how emotions and attendant thoughts – driven by *taṇhā* – *could* contribute to a sense of bounded self and (b) outlining Damasio's neurobiologically informed theory of self that lends strong independent support to this kind of account, bypassing the need to posit an independent bounded self in explaining how we come to have a sense of the self/other boundary.

2.2.1. How emotions and taṇhā might plausibly contribute to the impression of a bounded self

Any convincing account of self-as-construct must provide a plausible picture of how the sense of bounded self *could* arise through contribution from thought-processes, emotions and so forth, instead of from an actual unconstructed self – the entity from which it appears to arise. While providing such a picture will not alone prove the non-existence of an underlying self, it will be a first step in weakening the need to posit a *self* behind the sense of self. It should also clarify, on a psychological level, the potential role of various inputs into the sense of self, such as thought patterns, emotions and related dispositions.

There has been reason, so far, to suppose that emotion will play a significant role in the presentation of any self-sense and hence, in any potential account of the self-as-construct. For example, in Chapter 4 it was argued, in the section on 'Consistent Self-Concern', that the *presence* of most emotions (where there is apparent concern for a *self*) is enough to indicate a sense of bounded self (more on this soon). This argument was used to help justify the current chapter's contention that cases of depersonalisation are *not* good examples of awareness *sans* self-sense. And we reviewed evidence, in this current chapter, to suppose that a *lack* of emotion correlates with an absent sense of self, as exemplified in cases of epileptic automatism. The picture about to be presented will upstage the role of emotions from correlation to cause. It will be a role in which emotions, when tied up with *taṇhā*, are plausibly seen to both sufficiently indicate and *causally contribute* to the sense of a self-other boundary, such that an actual bounded self need not be brought in to explain the sense of bounded self. The picture is derived from my analysis of Buddhism and – it will later be seen – inspired by elements of Damasio's theory.

It is important, on this conjectured picture, to treat emotion as not merely a bunch of isolated affects (feelings), but as a cognitive state that incorporates those affects. As cognitive states, emotions will involve, as part of their ontology, a train of thoughts, sensations and perceptions with content usually depicting what the emotion is directed at, lending emotions their characteristic 'aboutness' or intentionality (although some moods might lack this). In Chapter 4 it was suggested that the content also reflexively

depicts – or seems to depict – a self who is having the emotion (this is also generally true of moods). Hence, on this picture, most emotions not only reflect a background desire to be happy or avoid suffering (a further feature of the self) but a background desire for one's *self* to be happy or to avoid suffering. Reflection on common emotions will support this picture. To feel guilty, for instance, usually implies that one wishes that one had performed an alternative action – one with a happier guilt-free outcome for *oneself*. To feel nervous at the thought of one's impending speech implies that one would like one's *own* speech to be well received and is concerned it may turn out badly for *oneself*, as opposed to any old person. One could not feel nervous, moreover, unless one identified as the future personal owner of the speech, believing it is the very same *self* as this current one who will bear the consequences.

Significantly, feeling nervous or guilty (for example) seems to indicate a *felt separation* between one's identified-as self and the imagined outcomes – whether desired or undesired. How could one feel nervous without at least implicitly cognising the possibility of the 'self's' alienation from success? Or guilty without believing the 'self' is regrettably *not* the owner of a more desired action? Such emotions thus seem to point to, as a part of their content, an unspoken *boundary* between the identified-as self on one hand and the desired or undesired scenario on the other, as it is perceived or imagined by the witnessing subject. From this picture, we can begin to see how such emotions might help *construct* the bounded self, such that the entity we assume we are exists as part of the *content* of the 'self'-concerned emotions.

However, while the content of such obviously felt emotions as nervousness, guilt and pride clearly exhibit and possibly contribute to the sense of a self-other boundary, it is also true that, throughout a given day, the attention of an average person is not likely to be grabbed all the time by the presence of such emotions. Many portions of waking life tend, in fact, to be rather emotionally neutral or fleeting, with a background mood whose tone might vary from motivated to bored. During such relatively neutral episodes, we can grant that the sense of ourselves in relation to perceived objects (implying a feeling of boundedness) will not be as obvious as it will be in the presence of stronger emotions. Nevertheless, the sense of self does not seem to *disappear* during such episodes of relative neutrality. If it did disappear, then we would expect the experience of perception-while-feeling-emotionally-neutral to significantly differ from that of perception-while-feeling-emotionally-charged – perhaps we should regularly feel depersonalised! – yet the situation does not present as being this way. Objects of perception seem presented to one and the same underlying self, emotionally charged or not. They all seem to be equally *my* perceptions. From my first-person perspective it seems to be exactly the same 'me' looking neutrally now at a blade of grass, as the 'me' who felt anxious earlier today. Now if emotion is supposed to be necessary for maintaining a sense of bounded self, how then do we explain such tracts of time

when there seems to be an ongoing sense of self without obviously felt emotion?

The Buddhist notion of *taṇhā* becomes relevant at this point. We can note, first of all, that *taṇhā* is the common factor to unite all those emotions that seem to depict the *self* as their protagonist. For in all such emotions, there is the background desire *for the self*, it would seem, to avoid suffering or become happier. *Taṇhā* is a mental disposition to want the world to conform to one's desires – one's *own* desires in particular – whether the desires are bodily or mental in their origin. The disposition involves emotional investment in the satisfaction of a desire, such that one's emotional state is affected by whether the desire is fulfilled. When the desire is not fulfilled, there is mental suffering in the form of a negative emotion such as mild annoyance, disappointment or anger. Conversely, when the desire is fulfilled, there can be such positive emotions as achievement, relief and joy.

It is plausible to suppose, then, that the majority of human emotions will involve some degree of *taṇhā*. But a mindset with *taṇhā* does not imply the obvious exhibition of emotion. If at any waking moment the affective tone of one's mindset, be this positive, negative or neutral, is dependent upon changeable circumstance, then whether one knows this or not, there is *taṇhā* present, according to Buddhism. We can for instance know whether there is *taṇhā* in our mindset if we sincerely answer 'yes' to the question: 'at a given waking moment, would any imaginable circumstance elicit happiness or mental suffering?' Most people, of course, would answer in the affirmative.

How could *taṇhā* actively create or maintain the sense of a self-other boundary in the absence of obviously felt positive or negative emotion? In other words, what causal effect could the very presence of *taṇhā* have upon one's thoughts and behaviour such that it propels an urge to seek one's 'own' happiness or avoid one's 'own' suffering (thus creating a persistent sense of self-other boundary)? I think it could work in the following way. It is not unreasonable to suppose that the very *thought* of experiencing a pleasant or unpleasant emotion – a thought that could occur in peripheral (or 'unprojected') consciousness – could exert a guiding effect on one's further thought and behaviour (and we will soon see a similar idea being supported in Damasio's theory). In this way, the Buddhist notion of *taṇhā* – the urge for 'ourselves' to be happy and avoid suffering – could conceivably drive much of one's thoughts and behaviour without the person being attentively cognisant of what is going on.

An example will help. We frequently change physical position; it is something we do with minimal cognisance. This action may well be motivated at least partly by *taṇhā*, if the unfulfilment of the physical desire to change position were to result in not only physical but mental suffering (as it usually would). The mental suffering, were it to manifest, would involve cognitive content implying a desire for the 'self's' separation from the physical

pain (e.g., the thought 'I wish it were more comfortable for *me*'), implying, in turn, a sense of self-other boundedness. Given this, it may well be the case that with most people, the urge to change position will be motivated not only by the instinctive desire to avoid physical pain, but also by the more cognitively loaded desire to avoid discomfort to *oneself*. The disposition to feel mental suffering at *not* changing position would be integral to this desire, and thus causally implicated in the actions taken to change position – even if we are not attentively aware of this desire or the actions caused by it. Given that it is 'my' suffering that is sought to be avoided (where there is a sense of 'mine' there must be a sense of 'me'), a sense of self-other boundedness would be part and parcel of this cognitively loaded desire. The mental discomfort almost unconsciously avoided in something as simple as changing position may thus involve a mental presentation of self-other boundedness that is likely to escape our attentive notice.

Should this Buddhist-based analysis be correct – and it seems on the face of it plausible enough – then it would be reasonable to suppose that during our waking lives, a sense of self-other boundary will almost continually be activated as we seek – attentively and inattentively – to bridge a gap between our current situation of relative or potential discomfort, and the one we imagine will reduce suffering or make us happier. As we go about our life, there will thus be an ongoing sense of *ourselves* as separate, bounded entities, in perceived or imagined relation to the desired or undesired situation. The cognitive medium for such perceivings and imaginings will be the stream of emotions and thoughts that imply *taṇhā*, including those that are subtle, imagined or unattended. According to Buddhism, this ongoing *taṇhā*, viz., emotional investment in desire satisfaction, will not only *indicate* but also help *create and drive* the sense of self-other boundedness. The sense of self-other boundedness, in turn, will fuel *taṇhā*. For unless I identified fully as a self, then how could I care particularly about whether 'my' desires are fulfilled? *Taṇhā* and the sense of self-other boundedness are hence plausibly co-dependent, on this Buddhist-derived analysis.

Now we might ask how this account would handle one's interaction with parts of the environment that do not elicit any emotion – perceived or imagined. For example, I could not care less about whether this blade of grass exists or not; yet I still identify myself as something decisively separate from the blade of grass. It is reasonable to suppose that if the Buddhist theory is correct, then the ongoing mindset of *taṇhā*, involving a constant (although not exclusive) lookout for 'number one', will be enough to create a general perceptual *tropism*, such that we involuntarily view even emotionally neutral items in our perceptual purview as decisively 'other' from the 'self' that perceives them.[7] *Taṇhā* will thus, on the Buddhist position, fuel the subject's overall disposition to *identify* itself with the perspectival owner of its perceptions, such that the perceptions are felt as personally 'belonging' to the

subject. In this way, then, we can see how the sense of self-other boundedness could be maintained without any need to posit an independent self.

On this conjectured, Buddhist-derived picture, then, the *taṇhā*-driven emotions, whether actual or imagined, attended to or not, are both sufficient indicators and causal contributors to the sense of a self-other boundary. Such a picture, which gives *taṇhā* a central role in constructing the sense of self-other boundedness, is consistent with those cases of epileptic automatism that were described as lacking in a sense of self. For Damasio's description of these cases not only pointed to a lack of obvious emotion (e.g., anxiety), but also to what seemed to be a complete lack of *taṇhā* or 'self'-concern. The patients exhibited no evidence of any background desire for their 'self' to be happy or avoid suffering, and hence, no evidence that they identified as the personal owner of their past, present or future experience.

2.2.2. Damasio's neurobiologically informed, parsimonious theory of boundedness as constructed from emotion, feelings and thoughts

We now consider some independent support for the Buddhist-derived analysis presented above. First of all, if the Buddhist analysis were correct in supposing emotional investment to be *necessary* for a sense of self, then we would expect to find no cases where *taṇhā* is absent but a sense of self present: no cases, that is, of a healthy sense of 'me' and 'mine' coupled with a lack of disposition to feel happiness or suffering in relation to *any* perceived or imagined states of affairs. As far as I know, no such case *has* been discovered. In support of this contention, Damasio (1999, 100) has concluded from his findings on epileptic automatism and other such pathologies that 'emotions [and feelings of them] and core consciousness tend to go together, in the literal sense, by being present together and absent together'. 'Core consciousness' on Damasio's (1999, 16) theory involves awareness coupled with a basic sense of self – although he never alludes to the 'awareness' component as a separate factor.

However, a realist theory of self is still not off the cards. Everything said so far is compatible with the idea that *taṇhā*, rather than helping construct the sense of self, is necessary for the reflexive *representation* of an independent, bounded self. In this story, a lack of *taṇhā* would simply mean that the self, while still existing, is not being reflexively represented. We would indeed *expect*, in this scenario, to find *taṇhā* co-existing with a sense of self. Cases such as epileptic automatism would simply mean that the usual faculties for letting the self be reflexively known to itself are suspended – just as they are suspended during deep sleep. In such instances the self would not cease to exist just because the *sense* of self is interrupted through a lack of wakefulness or emotional investment.

The time has finally come to dethrone this spectre of a realist theory of self. Damasio's constructivist theory is well suited to the task. As with the Buddhist account (itself partly influenced by Damasio's theory), it proceeds

by giving a convincing story of the apparent boundedness of self in a way that requires absolutely *no* reference to an unconstructed, bounded self which Damasio (1999, 11) sometimes terms a 'homunculus'. As with the Buddhist account, a mindset of *taṇhā* – which Damasio (1999, 304) refers to as 'an individual concern which permeates all aspects of thought processing, focuses all problem-solving activities, and inspires the ensuing solutions' – is considered central to the active generation and maintenance of the sense of self-other boundedness. The emotions are the vehicles through which this occurs. In contrast to the Buddhist account, however, Damasio buttresses the theory's psychological dimension with neurobiological support, usually from his own laboratory (although for ease of readership, my synopsis of his account will leave out the strictly technical aspects). From this perspective, he offers a persuasive theory of why a sense of self-other boundedness might have arisen in the organism's evolutionary history. Later, it will be seen how advantages of the Buddhist account in explaining *other* alleged features of the self (unity, elusiveness, etc.) can be used to strengthen what turn out to be weaknesses in Damasio's overall theory. The final account will thus draw upon philosophical and scientific resources that avert the need for having to introduce a special separate self-entity; it will explain the sense of self in a way that needs no recourse to a self.

The sense of self, on Damasio's theory, is basically a mental analogue of the organism's physical boundaries, which grows out of a biological need to maintain them. (This biological need, we soon see, is played out in emotions.) Hence,

> I have come to conclude that the organism, as represented inside its own brain, is a likely biological forerunner for what eventually becomes the elusive sense of self. The deep roots for the [sense of] self, including the elaborate [sense of] self which encompasses identity and personhood, are to be found in the ensemble of brain devices which continuously and *nonconsciously* maintain the body state within the narrow range and relative stability required for survival.
>
> (Damasio, 1999, 22)

This ensemble of brain devices, which Damasio calls the 'proto-self', is non-conscious. It is the neural co-ordination centre for the organism's homeostatic regulation. The sense of bounded self arises when, for reasons of biological advantage, the proto-self is *represented*, which involves the organism becoming conscious not only of the world around it, but of itself as an elusive entity interacting with the world. The sense of the self's individual boundedness and stability has its origins in the genuine but relative stability of the represented proto-self as it maintains homeostatic regulation within the organism's physical boundaries (Damasio, 1999, 168–194). However, I would explicitly add that the conscious representation of the

proto-self *as* a self is not veridical, but involves a cognitive exaggeration of its boundedness and stability as well as an additional sense of its being a conscious entity that is an owner, observer and agent. Were it not for this exaggeration and embellishment, the self would effectively *be* the biological proto-self that is represented – rather than the mental construct Damasio claims it is. It is important to bear this point in mind, which Damasio himself has not made clear enough – it will be returned to later.

The role of the emotions, on Damasio's theory, becomes relevant upon considering biological advantages that may be procured through representation of the proto-self as a self. While Damasio refers to emotions by familiar names and categories (e.g., the six 'primary' emotions, including happiness, anger, sadness, the social 'secondary' emotions such as jealousy, embarrassment, pride, and the 'background' emotions such as excitement, calmness, depression (1999, 51)), his use of the term 'emotion' is unconventional. For Damasio, 'emotion' does not denote the familiar first-personal feeling of these emotion categories; for this he reserves the term 'feelings'.[8] 'Emotion' denotes instead the 'complicated collections of chemical and neural responses' that have a regulatory role to play in helping the organism to maintain its boundaries (Damasio, 1999, 51). They are part and parcel of homeostatic regulation and can be thought of as 'sandwiched' between the most basic regulatory mechanisms, such as reflexes, and the 'devices of high reason' (Damasio, 1999, 54). The neurology of emotions will hence overlap with the neurology of the proto-self (Damasio, 1999, 100).[9] The most basic function of emotions is to assist the organism in maintaining its boundaries by avoiding sources of danger as well as pursuing sources of energy, sex or shelter (Damasio, 1999, 54). Now the effectiveness of emotions will be obviously improved, he suggests, if at least some of them become felt (from the first-person perspective) in a way that (a) connects them, through conditioning, to the fundamental motivating drives of pleasure and pain and (b) associates these feelings, through conditioning, with objects in the environment (Damasio, 1999, 54–55). In fact, I think that Damasio would hold that all emotions must at least have the *capacity* to be felt. Hence he writes:

> As a result of powerful learning mechanisms such as conditioning, emotions of all shades help connect homeostatic regulation and survival 'values' to numerous events and objects .... Emotions are inseparable from the idea of reward or punishment, of pleasure or pain, of approach or withdrawal, of personal advantage and disadvantage ...
>
> (1999, 54–55)

As the organism evolves in its complexity, its physical boundaries will become increasingly vulnerable, requiring devices more ingenious than blind conditioning to protect them. This is where foresight – requiring consciousness with a sense of self – comes in. The capacity to plan ahead or

represent past scenarios means the organism need not wait until danger is immanent or the energy source proximate, in order for it to act accordingly. Foresight, reason and extended memory enable the organism to mentally review its options for action and response – and then choose what is most beneficial to it. The capacity of the organism to represent itself (viz., organism) *as* a past, present or future *self*, is, on Damasio's theory, an adaptation that grows out of the organism's need to foresee and plan its future in an optimal way (Damasio, 1999, 139–140, 284–285, 303–304). Reviewing one's future becomes viscerally motivating when infused with the sense that one's *own* future is at stake.

Now this process of mental reviewing and choosing, according to Damasio, is not an exercise of pure reason, but is very much grounded in the emotions which have, through conditioning and memory, become associated with certain options. Let us take an (rather effete!) example. Suppose you feel like socialising with friends (an option you have habitually come to enjoy) but have a lot of work to complete by the next day. Distinct feelings of anxiety (associated with pain) arise, alongside the thought 'Unless I get this work done today I may lose the job I like!' The emotions that arise in connection with the imagined future scenarios are, of course, not felt to be anyone's old emotion – they are felt to be *your* emotions in connection with *your* possible future. In weighing up the options you seem to be imagining *yourself* going out with friends or working, bearing the emotional consequences of each. The emotions that are felt as yours, whether imagined or actual, help you to decide to stay at work.

So as well as allowing the organism to represent itself as a past or future entity, consciousness with a sense of self, Damasio (1999, 81) holds, allows the organism's key guiding emotions, in the form of feelings, to be immediately known to it. This enables the organism to feel more of an incentive to act on them – as it will need to do when imagining past or future scenarios. And the organism's emotions (as feelings) cannot be known to it – at least in this extra-motivating manner – unless it represents *itself* as the elusive and *personally concerned* bearer of the emotional feelings, hence, as a distinct self (Damasio, 1999, 285, 302–304). So whenever emotions, or for that matter any other mental or perceptual states are felt to be 'mine', the proto-self (sharing neurology with emotions) is being represented as a 'me', viz., as a self who has – and can act in accordance with – these emotions and perceptions, etc.

Now we saw that what compels the organism to be initially motivated by the feelings of emotion (and hence to protect its boundaries) are the conditioned associations of the emotions with pleasure and pain. This initial motivation (and effectiveness at boundary-preservation) is greatly enhanced by a feeling of overriding and asymmetrical *concern* that is bestowed by the sense of self (which Buddhism calls '*taṇhā*'). The 'self' feels, in the words of William James (1890, 297), 'the home of interest, – not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and

pain, the pleasant and the painful, speak'. This overriding concern will ensure that the organism-subject (identifying as a *self*) is primarily motivated by *its* pleasure and pain: looking out for number one, so as to protect *its* boundaries – both psychological and biological. Hence, Damasio writes:

> I would say that consciousness, as currently designed, constrains the world of imagination to be first and foremost about the individual, about an individual organism, about the self in the broad sense of the term … Perhaps the secret behind the efficacy of consciousness is selfness. In short, the power of consciousness comes from the effective connection it establishes between the biological machinery of individual life-regulation [the proto-self] and the biological machinery of thought. That connection is the basis for the creation of an individual concern which permeates all aspects of thought processing, focuses all problem-solving activities, and inspires the ensuing solutions. Consciousness is valuable because it centers knowledge on the life of an individual organism.
>
> (1999, 304)

To summarise, on Damasio's theory, the sense of self, integral to 'core consciousness', grows out of a biologically driven need for feelings of emotions to be known to the organism having them. Knowing emotions – which requires thought and feeling – enables the organism to plan effectively for its future, by envisaging itself as the past and future subject of positive or negative scenarios. The ongoing motivation to avoid the harmful and pursue the beneficial is made possible by an ongoing individual concern that permeates the organism's psyche, a concern which is felt through the psychological medium of these thoughts and feelings. Biologically, this concern is grounded in the innate drive for homeostatic boundary-preservation; psychologically, this concern manifests as a sense of wanting the best for one's *self*. On a psychological level, the concern effectively creates the sense of self-other boundedness and the sense of boundedness in turn, perpetuates the concern. As with the Buddhist account, *taṇhā* and a sense of the self's boundaries are co-dependent.

Now, at any waking moment, one's impression of the self as owner, thinker and actor will, on Damasio's theory, be generated by a stream of perceptions, thoughts and attendant emotions which represent not only the objects attended to, but the organism that is attending to them. It represents the organism *via* this impression of a self that engages with the objects. And in keeping with the Buddhist-derived account, the cognitively construed emotions (with thoughts and feelings) that subtend *taṇhā* and allow for a continuing sense of self are abundant, with many of them relatively subtle, imagined and unattended to:

> In effect, normal human behaviour exhibits a continuity of emotions induced by a continuity of thoughts. The contents of those thoughts, and

> they are usually parallel and simultaneous contents, include objects with which the organism is actually engaged or objects recalled from memory as well as feelings of the emotions that have just occurred. In turn, many of these 'streams' of thought – of actual objects, of recalled objects, and of feelings – can induce emotions, from background to secondary, with or without our cognizance. The continuous exhibition of emotion derives from this overabundance of inducers, known and not known, simple and not so simple.
>
> (Damasio, 1999, 93)

It is as if, Damasio contends, the thoughts and felt emotions are both verbally and non-verbally telling one that they pertain to *me*, a *self*, and that *I* had better heed them to maximise my welfare. Damasio uses the term 'images' to convey the cognitive medium by which the thoughts and emotions are consciously known to one. (A more familiar word might be *qualia*.) Hence the thoughts and emotions, through mental imagery, tell the organism a story of itself, except that in this story, the organism does not feature as merely a biological entity, but as a *self* engaging with the world (Damasio, 1999, 188–192). While the self is depicted in these images as the story's protagonist, there is no *real* self (or homunculus, as Damasio puts it) who is telling the story. The self, as such, is an illusion. Hence,

> You know you exist because the narrative exhibits you a protagonist in the act of knowing. (Damasio, 1999, 172) … The story contained in the images of core consciousness is not [however] told by some clever homunculus. Nor is the story really told by *you* as a self because the core *you* is only born as the story is told, *within the story itself*. You exist as a mental being when primordial stories are being told, and only then … You are the music while the music lasts (Damasio, 1999, 191).

The constructed, hence illusory status of self is thus made plain on Damasio's account, although, perhaps because he has not defined the notion of an illusion, Damasio never *says* that the self is an illusion.

I hope to have conveyed at least some of what is appealing about Damasio's position (naturally I will not have done it full justice in these few paragraphs). In my opinion, his theory offers powerful reason to suppose that positing a thought-independent bounded self to explain our impression of the self's boundedness is superfluous – both diachronically and synchronically. From an evolutionary (diachronic) perspective, Damasio provides a convincing (although in places speculative) account of how the sense of personal ownership and hence boundedness might arise from adaptational pressure on the organism to plan for its future with the aid of emotional feedback. Synchronically, he provides convincing reason to suppose that the bounded self arises, at a time, through the content of felt emotions

and thoughts which 'tell the story' of such a self who is having them. (Evidence: knock out emotion and the sense of self disappears.) The self-referencing thoughts and emotions, driven by concern for one's 'own' welfare, are plausibly ubiquitous enough to allow for such a sense of bounded self to be sustained. Owing to conditioning which associates emotional reaction with numerous objects, internal or external, the relevant triggers for emotion and thought will be profuse, including the self-referencing thoughts and emotions that will themselves become triggers.

Damasio's theory, then, gives independent credibility to the Buddhist proposal that the self is a construct whose existence and continuation as a *personal, bounded entity* psychologically depends upon *taṇhā*, that pervasive sense of concern for one's 'own' welfare *qua* self. This *taṇhā*, both driving and borne out by a retinue of (often subtle, unattended to) felt emotion and thought, serves also to protect the organism's *physical* boundaries (e.g., by making decisions to avoid immanent danger). By virtue of its boundedness, the self will therefore be a construct: the content of its appearance contributed to by the content of those thoughts and emotions that subtend the *taṇhā*. Now if the self is a construct then it must also be illusory. For the self, as we know, purports to be an unconstructed entity that *has* the thoughts, emotions and perceptions rather than something that is any way created by them.

2.3. The shortfalls of Damasio's theory

While Damasio has, I believe, provided a plausible theory of how the self comes to exhibit boundedness, there are gaps in his account of the self as a whole. Indeed, Damasio does not focus enough on how the representation of the proto-self as a self ends up exaggerated or embellished with *other central features* attributed to the self. So while stable, relatively speaking, the proto-self is not unbroken or invariable in the uncompromising manner that the self purports to be. While its neurons work together to serve the individual organism, the proto-self is not unified in the same uncrackable way that the self purports to be (in never seeming, to itself, to be disunified). Nor are we even conscious of the proto-self – and it is definitely not conscious of itself – a point that Damasio (1999, 174–175) freely admits. Lacking in consciousness, the proto-self can hardly, therefore, be elusive to its own attentive observation (coupled with a subtle sense of its own existence). Hence, as a purportedly conscious entity, the self displays features that are either exaggerated or not belonging at all to the represented proto-self. While Damasio does attempt to deal with elusiveness (1999, 172) and to some extent unbrokenness (1999, 176), his account of how these features arise as integrated with the bounded self is unconvincing. If my arguments so far are correct, then his account of these features must fall short, since its very framework ignores explicit reference to *awareness*, the immediate ground of these aforementioned features. Being intrinsic to awareness, these

features cannot be entirely constructed, as Damasio proclaims. When he talks about 'the knower' or 'consciousness', moreover, it is always in the context of awareness with a sense of self. His theory does not cater for awareness or witnessing *simpliciter*.[10]

In what follows, I propose to extend a main aspect of Damasio's theory in a direction that I think will accommodate awareness and hence account properly for some main features, besides boundedness, that are ascribed to the self. The proposed account will not be offered at the neurologically informed level of Damasio's theory; it will be proposed at a 'higher' psychological level. There is a reason for being optimistic that the grafting will work: the Buddhist theory of self-other boundedness, which links the sense of boundedness to *taṇhā*, finds much support in Damasio.

The part of Damasio's theory that I propose to extend, then, is the aspect where the sense of bounded self arises from the *taṇhā*-driven stream of thoughts and emotions that seem, first-personally, to imply the existence of an *individual, bounded self* as their thought-independent bearer. I propose to do this by first reviewing the concept of identification (proposed in Chapter 4) in light of the self being a construct and illusion. I then integrate it with the relevant aspects of Damasio's account.

2.4. Identification revisited in light of self as illusion

Identification serves, as we saw in Chapter 4, to separate a sense of mere *perspectival* (or possessive) ownership from *personal* ownership. For an item, X, to be (and seem) perspectivally owned by subject S is for X to appear to a subject S in a way that it appears to no other subject. Such Xs typically include thoughts, volitions, emotions, sensations and perceptions. Since these cue a subject into viewing their body and actions from a unique perspective, the body and actions can also be regarded, more loosely, to be perspectivally owned. The perspectival owner of an X is effectively a subject: awareness as it appears from a body-bound, sensory-mental perspective in relation to X. Note that there is nothing illusory about the perspectival owner itself; it involves merely the contingent and undisputable fact of awareness normally appearing from the psycho-physical perspective of a body and mind.

For something perspectivally owned by S to seem also *personally* owned by S is, minimally, for S to *identify* with the perspectival owner of that item, X. For S to identify with the perspectival owner of X is for S to appropriate to its (witnessing) perspective those psycho-physical cues that lend the subject a distinct perspective in relation to X – often reinforced by an *idea* (such as a memory) of 'being' X's perspectival owner. Through appropriation, the items seem to become infused *with* the witnessing-from-a-perspective, such that they lose any object-like appearance. Whenever S's current witnessing perspective takes itself to *be* the perspectival owner of X, through appropriation, feelings of *personal* ownership will automatically arise towards X. The

familiar feelings of personal ownership or my-ness will seem in turn, to wordlessly indicate not just a perspectival but a personal owner who has X – a self, in other words. In this way we can see how the minimally contingent perspective of awareness – the perspectival owner – becomes folded into the sense of a personal self. (A similar account is given by which a subject can appropriate to its perspective the idea of being a *possessive* owner – often in relation to an assumed, prior identity.)

In Chapter 4, I left open the possibility that those feelings of personal ownership actually do indicate a personal owner or self who has them. I hence left open the possibility that S's current witnessing perspective might actually be part and parcel of a self (that awareness might be bounded, in other words). On such a possibility, S's identification with a perspectival owner of X would amount to identification with a *personal* owner of X; identification would hence double as the veridical and reflexive *representation* of a self. This would then imply that when we seem to recall ourselves (viz., a self) doing Y (involving apparent identification *with* personal owner of action Y) then we are, quite literally, recalling our *self* doing Y. Or when we imagine what we assume is our future self doing Z then we are literally imagining our *self* doing Z.

Now if the self is, as I have suggested it is, a construct, then a subject's identification with a perspectival owner, such as through memory, will not be coupled with the actual representation of a personal owner. Identification, which elicits feelings of personal ownership, will only *seem* to represent a personal self. So when we seem to recall *ourselves* performing an action, what will actually occur is that we appropriate, to our current witnessing perspectives, those psycho-physical cues (subtending the act of remembering) that stand our perspectives in relation to the remembered action. Such appropriation will elicit feelings of personal ownership towards the past action and those feelings will in turn seem to point to a *personal* owner who performed the action (and is now remembering it). But there will be no actual personal owner; such is the illusion of self.

An example will help. Suppose Ben wins a race, and then feels proud because of it. On our analysis, Ben's feeling of pride and attendant thoughts 'I won the race!' indicate that he feels personal ownership towards winning the race and his proud feelings. The feelings of pride are a give-away sign that, as a current witnessing perspective, Ben identifies *with* the perspectival owner and agent who won the race (appropriating those memory-cues to his current perspective) and hence *as* the personal owner and agent who won. He would not feel proud unless he believed it was *he* who won the race. But is this belief a correct one? Not literally. He believes that he *qua personal self* won the race – that is how the event seems remembered from his perspective. His believing this much, however, has already bought into the content of those thoughts and feelings that tell him 'I, a personal owner and agent called 'Ben', won the race, and I, the same personal self who won the race,

am feeling proud because of it'. But such feelings, I have argued, mislead. There is in fact no such self-entity who could have won the race or now believes that he won it. There is only, at any one time, a perspectival witnessing presence with a bundle of thoughts and emotions that, through the mechanism of appropriation, collectively comprise the *sense* of a self that owns them. There is no actual personal self who has these emotions; that is what it means to say the self is an illusion.

2.5. Integrating concept of identification with Damasio's analysis to yield the two-tiered illusion of self

We are now in a position to see how the proposed theory of identification could dovetail into the Damasio's analysis of self-other boundedness. On the proposed theory, the feelings of personal ownership of my-ness, underpinning the impression of a bounded self will, as Damasio suggests, have major contribution from a stream of often subtle thoughts and (felt) emotions. These thoughts and emotions will keep alive the current of *taṇhā*, their content conveying, among other things, a bounded self looking out for its *own* interests. It is *this* component of the personal-ownership feelings – the emotions and their attendant thoughts (presented through mental imagery, augmented with physical sensations) – that serve to *actively* construct the boundary between oneself and the rest of the world.[11] But – and here is where the proposed account expands on Damasio – the feelings of personal ownership that tell the story of a self will also involve input from the *current awareness*: the witnessing presence that *identifies as* the self who is the personal owner of the feelings. The native input from awareness, as it identifies with various items (and *as* a self in such a capacity), will serve as the immediate origin of many features ascribed to the self as a whole – those features that qualify awareness simpliciter. They will be the features that Damasio's account cannot properly explain: unity, elusiveness, unbrokenness, invariability, unconstructedness and, of course, awareness itself.

Along with the advantage of explaining major features ascribed to the self, the integration of Damasio's theory with the Buddhist-derived account of identification offers a further advantage. In Chapters 3 and 4, I spoke of some research benefits to empirically linking identification and the personal ownership-sense (and hence a sense of self) with *taṇhā*. Now that the link has proved viable, we have an acid test for determining not only whether a person harbours a sense of self, but a test for determining the *particular items* towards which one harbours a sense of self (viz., a sense of personal ownership and self-identification). Such items are most often (a) perspectivally or possessively owned *and* (b) have a perceived status that affects the actual or potential happiness or mental suffering of the perspectival/possessive owner. It can thus be confidently asserted that emotional investment (*taṇhā*) towards a perspectivally or possessively owned item is, empirically speaking, both a necessary *and* sufficient sign that one identifies as that

item's personal owner. Where one has *taṇhā* towards an item that is not perspectivally or possessively owned (such as a political situation or another person), one will harbour a sense of personal ownership and identification towards various thoughts, emotions and perceptions that *relate* to that item – including the sensory mediums through which that item is known to the subject. In all such cases, the presence or absence of *taṇhā* – and hence a sense of self – can be objectively tested for. One possible way to do this would be to measure galvanic skin conductances in response to stimuli that pertain to suspected items of *taṇhā* (a procedure that Damasio (1999, 49–50) has himself utilised in measuring emotional response to stimuli).

Damasio's account, furthermore, is peppered with evidence that the neurobiological correlates to the sense of bounded self will involve what he calls the 'proto-self' (e.g., Damasio (1999, 100)). And we may recall that Damasio provides evidence to suppose that the proto-self shares its neural substrate with the (non-conscious) emotions whose job it is to maintain homeostatic regulation (1999, 100). The significance of such a connection is to place the ontology of the self-sense in a neurobiological context, taking seriously the emotions of *taṇhā* on both a conscious and non-conscious level. This provides another potentially objective means by which one could test for the presence or absence of *taṇhā* – and hence for a sense of self. All this will have important implications for exploring, down the track, the credibility of the Buddhist claim that practices such as meditation gradually eliminate the sense of a bounded self (while retaining awareness).[12]

In the following and final chapter, I set out to more formally describe the model of the two-tiered self-illusion that has been argued for in this chapter. The two tiers consist of (1) unconstructed 'awareness' (qualified with its intrinsic features) and (2) those *taṇhā*-driven thoughts and emotions that create the impression that such awareness is part and parcel of a *bounded* self.

9
The Two-Tiered Illusion of Self

1. A schema for the two-tiered illusion of self

Up until now, I have described the illusion of self in terms of the broad schema from Chapter 5. Thus, I have been saying that the self-as-unconstructed is illusory; for while it purports to be unconstructed, it is, as a matter of fact, constructed. Given the more detailed analysis from Chapter 8, which pins the source of the self's constructed status on boundedness (and not awareness), I now propose a more specific way to describe the illusion of self. I call it the 'two-tiered illusion of self'. The two tiers in the self-illusion do not correspond to two ontologically separate processes that occur during identification. They correspond, rather, to two strands or 'feeders' that contribute to the single process of identification, a process that creates the illusion of a *self*. I call these two inputs into the self-illusion the 'boundedness feeder' and the 'awareness feeder'.

The first tier of the illusion consists of input from the 'boundedness feeder' and can be described as a person's awareness (with its co-specifying features), actually *non-bounded* in its nature, purporting to be a *bounded self*. Since awareness is not inherently bounded in this way, it cannot *be* the bounded self that it purports to be: hence the illusion of self, viz., awareness-as-bounded. The second tier of the illusion consists of input from the 'awareness feeder'. It can be described as the bounded self, actually a *construct*, purporting to harbour those features that rightly belong only to the *unconstructed awareness*. Since the self is not inherently unconstructed in the manner it purports to be, it cannot be an unconstructed self: hence the illusion of self, viz., self-as-unconstructed.

The idea behind this 'two-tiered' illusion of self is not, of course, original to this project; it has prefigured in Buddhism and, more explicitly, in Advaita Vedanta. In the following passage, the leading figure of this tradition, Śaṃkara, (788–820 CE), speaks of awareness (akin to undifferentiated fire or clay) taking itself to be a bounded entity (akin to heated iron or an

earthen jar), by adopting features associated specifically with this bounded entity (e.g., personal ownership):

> Allied to the intellect, just a part of itself, although the true self of everything, and beyond the limitations of such an existence, [awareness] identifies itself with this illusory self - as if clay were to identify itself with earthen jars.
>
> (Śaṃkara, 788–820/CE, Verse 190)

> In conjunction with such additional qualities, the supreme self [awareness] seems to manifest the same characteristics, just as the undifferentiated fire seems to take on the qualities of the iron it heats.
>
> (Śaṃkara, 788–820/CE, Verse 191)

And the bounded self through which awareness is reflected, appears, through identification, to take on what is inherent to awareness (akin to the sun reflected in a water-jar being mistaken for the sun):

> The ignorant see the reflection of the sun in the water of a jar and think it is the sun itself. In the same way [he] sees the reflection of consciousness [viz., awareness] in its associated qualities [to do with boundedness] and mistakenly identifies himself with it.
>
> (Śaṃkara, 788–820/CE, Verse 218)

The natural state of awareness is, however, unfettered by illusions of boundedness. To realise its true nature is to see through this illusion:

> The wise man ignores jar, water and the sun's reflection in it, and sees the self-illuminating sun itself which gives light to all three but is independent of them.
>
> (Śaṃkara, 788–820/CE, Verse 219)

I now elaborate in further detail on how to construe the two-tiered illusion of self such that we see how it is borne out through (a) co-specified awareness in general and (b) each feature intrinsic to awareness (which purports to be unconstructed) – witnessing presence, unity, elusiveness, unbrokenness and invariability. The most precise way to execute this task will be to use a template for 'Tier One' and 'Tier Two', applying it systematically to each feature. While some repetition and overlap will be inevitable, the exercise should serve to draw many threads of argument from previous chapters into the beginning of a unified, patterned weave. I say 'beginning' because it will be obvious that the proposed account of the self's ontology is sketchy and needs further development. Nevertheless, one must be reminded that other areas which might seem flimsy in this presentation are in fact

recapitulations of earlier argument – for example, the evidence that we identify as a *self* with each of the enlisted features, and the justification for supposing that the *self,* with such a feature, is illusory.

The following model of the two-tiered self-illusion thus applies to the co-specified awareness in general, and will serve as the template for examining in more detail how the schema applies to each specific feature of awareness, including its central core of witnessing presence.

The *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as 'awareness-*as-bounded*'. *Evidence* that we do indeed reflexively construe awareness to be *bounded* resides in the fact that feelings of personal ownership such as this-ness convey a *personal* as opposed to a merely *perspectival* subject who stands behind and owns these feelings (rather than seeming to be constructed by them). Our *explanation* for this apparent *boundedness* of awareness is that through the process of identification (driven by *taṇhā*) the 'story line' told by current emotions (obvious or subtle, attentive or inattentive), with attendant thoughts and perceptions, 'casts' the natural input of awareness as a bounded self. In this 'boundedness' tier, awareness-as-bounded comes out as *illusory* because awareness purports to be bounded when awareness is not in fact bounded in its nature. Awareness is not bounded because we have shown that (a) *awareness* in itself is unconstructed (from the arguments of Chapters 6 and 7) while (b) awareness-as-bounded is constructed, viz., the content of an appearance by which it is contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since awareness is unconstructed, it cannot be bounded (since awareness-as-*bounded* is a construct), which renders the self, viz., awareness-as-bounded, an illusion.

The *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded-awareness>, is presented as being as *unconstructed* in the manner of pure awareness. *Evidence* that we construe the self, viz., <awareness-as-bounded>, as *unconstructed* is that, from the viewpoint of 'myself', it never feels as if the self as a whole is something whose existence is dependent upon current emotions and their attendant thoughts and perceptions. The self is always depicted as their personal, ontologically unique, *owner*, standing apart from them as a separate entity. Our *explanation* of this impression is that in the process of identification, the real unconstructedness of native awareness filters into the net impression of the bounded self-entity. In this tier, bounded-awareness-as-unconstructed comes out as *illusory* because the bounded awareness, viz., the self, purports to be unconstructed when as a matter of fact the self is a construct The self is not actually unconstructed (as it purports to be) because, as demonstrated in Chapter 8, the boundedness aspect of self-as-unconstructed is the content of an appearance by which it is contributed to by perspectivally ownable objects, rather than existing independently of these objects as it

purports to do. Evidence for this constructedness is that a suspension of emotion and its attendant thoughts results in the suspension of the impression that the awareness is bounded (namely, a sense of self). Hence the self, viz., bounded-awareness-*as-unconstructed*, is an illusion.

I now apply this schema of the two-tiered self-illusion to awareness in light of each of its specific features (each of which purport to be unconstructed); first to its *raison d'être* of witnessing presence, and then to unity, elusiveness, unbrokenness and invariability. In the final two applications of the schema, we start to get a picture of how the impression of *longer-term personal identity* may arise from unbroken and invariable awareness feeding into the content of (*taṇhā*-driven) thought and emotion through the process of identification.

2. Witnessing presence

With regard to witnessing presence, the *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as '*witnessing-presence*-(awareness)-*as-bounded*'. *Evidence* that we do indeed reflexively construe witnessing presence to be bounded resides in the fact that feelings of this-ness and dispositional value seem to point to 'this very presence who witnesses'. In other words, witnessing-presence does not normally present as something that is non-bounded and impersonal. Our *explanation* for this apparently personal boundedness of witnessing presence is that, through the process of *taṇhā*-driven identification, the 'story line' told by current emotions (obvious or subtle, attentive or inattentive) with attendant thoughts and perceptions 'casts' the natural input of witnessing presence as a bounded self, perhaps in a similar way that dreams can 'cast' real alarm sounds into their story line as shrill voices. In this boundedness tier, the self, viz., *witnessing-presence-as-bounded*, comes out as *illusory* because *witnessing-presence* purports to be *bounded* when *witnessing-presence* is not in fact bounded in its nature. *Witnessing-presence* is not bounded in its nature because we have shown that (a) *witnessing presence* in itself *is* unconstructed (from the arguments of Chapters 6 and 7) while (b) *witnessing-presence-as-bounded* is constructed, viz., the content of an appearance by which it is contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since *witnessing presence* is unconstructed, it cannot be bounded (since *witnessing-presence-as-bounded* is a construct) which renders the self, viz., *witnessing-presence-as-bounded*, an illusion.

With regard to witnessing presence, the *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded-witnessing-presence>, is presented as being as *unconstructed* in the manner of pure awareness. The *evidence* that we construe the self, viz., <bounded-witnessing-presence>, as *unconstructed* is that from the viewpoint of 'myself',

it would seem that whatever it is that makes my witnessing-presence seem uniquely and *personally* mine will be unobservable as an object – (it cannot be attended to) – thus apparently excluding objects from any role of constructing the boundaries of this witnessing perspective. Our *explanation* of this impression is one that appeals to genuine unconstructed input from the awareness feeder such that an impression of an *unconstructed* bounded witnessing presence is generated. Whenever a current witnessing presence appropriates an idea to its perspective, the native contribution from awareness to resultant feelings of personal ownership will naturally infuse 'the personal owner' with that subjective sense of presence. The witnessing presence is what lends the apparent self its first-person, subjective perspective. Because a 'personal witnessing presence' is all that most of us have experienced (at least to the best of our memory), we normally cannot imaginatively separate pure first-person feelings from *personal* first-person feelings – just as a saltwater fish (should it have the relevant conscious capacities!) could not imaginatively separate the sensations of freshwater from that of the saltwater to which it is accustomed. To the usual conscious mind, witness-presence awareness and 'me' thus present as a basic unit.

In this 'awareness' tier, the self, viz., *bounded-witnessing-presence*-(awareness)-*as-unconstructed* comes out as *illusory* because the *bounded-witnessing-presence* purports to be unconstructed when as a matter of fact, as demonstrated in Chapter 8, it is a construct. The boundedness aspect of witnessing-presence-as-unconstructed is the content of an appearance by which it is contributed to by perspectivally ownable objects that subtend *taṇhā*, in particular, thoughts and emotions. Evidence for this constructedness is that a suspension of emotion and its attendant thoughts results in the suspension of the impression that the witnessing perspective is bounded (through the sense of being a personal owner). Hence the self, viz., bounded-witnessing-presence-*as-unconstructed*, is an illusion.

3. Unity

With regard to unity, the *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as '*unified*-awareness-*as-bounded*'. *Evidence* that we do indeed reflexively construe unified awareness to be *bounded* resides in the fact that any observable thoughts and impressions at a given waking moment will seem to present themselves not merely to a unified awareness but to a unified *personal* awareness. The thoughts will hence not merely seem to be perspectivally mine, but personally mine, indicating that I identify as a *personal unified owner* who has them. The self as a bounded entity is felt to *be* that entity which unites such items as its (personal) own. And the entity of self will seem to radiate a unity that extends not just to awareness, but to the complex traits of an entire personality. Our *explanation* for this apparent boundedness of unified awareness is

that, through the process of *taṇhā*-driven identification, the 'story line' told by current emotions (obvious or subtle, attentive or inattentive) with attendant thoughts, feelings and perceptions 'casts' the natural input of unified awareness as a bounded self. The story line is replete with references to a *personality* which, when identified with as a self, will seem to mold the unity of witnessing to a personal shape. In this boundedness tier, *unified*-awareness-*as-bounded* comes out as *illusory* because unified awareness purports to be bounded, when in fact unified awareness is not bounded in its nature. Unified awareness is not bounded because (a) *unified awareness* in itself is unconstructed (from the arguments of Chapters 6 and 7) while (b) *unified*-awareness-*as-bounded* is constructed, viz., the content of an appearance by which it is contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since *unified*-awareness is not constructed in actuality, it cannot be bounded (since unified-awareness-as-*bounded* is a construct) which renders the self, viz., *unified*-awareness-*as-bounded*, an illusion.

With regard to unified awareness, the *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded unified awareness>, is presented as being *unconstructed* in the manner of pure awareness. The *evidence* that we construe the self, viz., <bounded unified awareness> as *unconstructed*, is that it would seem to 'me' that 'I', a unified self to whom impressions belong, am *necessarily* unified, with no impression of any potential cracks in this unity (an impression that would be present if such unity *seemed* to overtly depend upon a flux of thought and emotion). As a result, it would seem to me that I could not really imagine what it would be like for my personality-infused *self* (not merely my consciousness) to lack the impression of unity. Our *explanation* of this impression will be that during the process of identification, the mechanism of appropriation *somehow* employs the native unity of awareness to 'fuse itself together' with any appropriated thoughts or ideas, such that there is the impression of an organised locus for a unified personality. The feelings of personal ownership will thus seem to unquestioningly emanate from a personal owner who *is* this necessarily unified entity, presiding over the thoughts that 'it' thinks and owns. In this tier, the self, viz., *bounded-unified*-awareness-*as-unconstructed*, comes out as *illusory* because the *bounded-unified*-awareness purports to be unconstructed when, as a matter of fact it is, as demonstrated in Chapter 8, a construct. The boundedness aspect of unified-awareness-as-unconstructed is the content of an appearance by which it is contributed to by perspectivally ownable objects that subtend *taṇhā*, in particular, thoughts and emotions. This is evidenced in the fact that the apparent unity of self can be shown to lack the necessity that it subjectively seems to have. Cases of epileptic automatism (with their suspension of thought and emotion) together with our background conclusion that the bounded self is a construct, are

testimony to the fact that the 'unity of self' can come apart. Hence the self, viz., bounded-unified-awareness-*as-unconstructed*, is an illusion.

4. Elusiveness

With regard to elusiveness, the *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as '*elusive*-awareness-*as-bounded*'. *Evidence* that we do indeed reflexively construe elusive awareness to be *bounded* resides in the fact that it is not only awareness that seems elusive to its own observation but *personal*, *bounded* awareness. Hence *boundaries* seem part and parcel of what is elusive about awareness. While having a sense of its 'own' boundedness, it is a seemingly *personal* awareness that can never manage to observe its boundaries as an object of thought or perception. The boundaries appear to qualify an elusive, observing, personal *subject* – as opposed to a non-elusive, observable *object*. Our *explanation* for this apparent boundedness of elusive awareness lies in the stream of thoughts and emotions that (through *taṇhā*) contribute a sense of personal ownership or my-ness towards any object of consciousness. While the lion's share of the 'aboutness' of these thoughts and emotions is the main target of the story line, (e.g., our impression of the objects to which the thoughts and emotions seem directed), a portion of this 'aboutness' is reserved for the implicit self who is having them, viz., the personal owner of the thoughts and emotions. This sidelined section of the story line's 'aboutness' always portrays the self as an individual protagonist who is definitively *separate* from the target object, helping to create a distinct impression of self-other boundedness.

In this 'boundedness' tier, the self, viz., *elusive*-awareness-*as-bounded*, comes out as *illusory* because elusive awareness purports to be bounded, when in fact elusive awareness is not bounded in its nature. Elusive awareness is not bounded because (a) *elusive awareness* in itself is unconstructed (from the arguments of Chapters 6 and 7) while (b) *elusive*-awareness-*as-bounded* is constructed, viz., the content of an appearance contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since *elusive* awareness is not constructed in actuality, it cannot be bounded (since elusive-awareness-as-*bounded* is a construct) which renders the self, viz., *elusive*-awareness-*as-bounded*, an illusion.

With regard to elusive awareness, the *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded elusive awareness>, is presented as being *unconstructed* in the manner of pure awareness. The *evidence* that we construe the self, viz., <bounded elusive awareness> as *unconstructed* lies in the fact that it does not appear as if the 'bounded awareness' can ever observe its 'own' boundaries as a non-elusive object – which it would need to be able to do if the bounded elusive awareness were to *seem* constructed by perspectivally ownable objects. Our *explanation*

of this impression is that the elusiveness, which seems to make the bounded self-entity sidelined and unobservable, will in fact emanate from the elusive awareness that has contributed to the story line of its 'protagonist'. Through the mechanism of appropriation, awareness will somehow be blinded to the fact that 'its elusive boundaries' are nothing more than the projected and imagined content of thoughts and emotions, together with its own native input of elusiveness.

In this tier, the self, viz., *bounded-elusive*-awareness-*as-unconstructed*, comes out as *illusory* because the *bounded-elusive*-awareness purports to be unconstructed when as a matter of fact it is, as demonstrated in Chapter 8, a construct. The boundedness aspect of elusive-awareness-as-unconstructed is the content of an appearance that is contributed to by perspectivally ownable objects that subtend *taṇhā*, in particular, thoughts and emotions. There are hence no elusive boundaries of self, since there is no self that is bounded. There is only elusive awareness together with a bundle of non-elusive thoughts, emotions and ideas that, with the help of elusive awareness, tell the story of an elusive self. Hence, the self, viz., bounded-elusive-awareness-*as-unconstructed*, is an illusion.

5. Unbrokenness

With regard to unbrokenness, the *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as '*unbroken*-awareness-*as-bounded*' from moment-to-conscious-moment. And while there is no obvious impression of unbroken awareness during episodes of deep sleep, there is nevertheless an impression that the self has numerical identity over longer spans of time. I suggest that the natural unbrokenness of awareness (from one conscious moment to the next) feeds into both impressions. *Evidence* that we do indeed reflexively construe unbroken (moment-to-moment) awareness to be *bounded* lies in the fact that this unbroken element of our consciousness life seems unified, elusive and *personal*. On a longer-term basis, it seems to be numerically one and the same bounded personal self that is living this life with a history and anticipated future.

Our *explanation* for the *apparent boundedness* of unbroken awareness is as follows. Regarding the appearance of *bounded, unbroken awareness*, the constant stream of thoughts and emotions (etc.) being appropriated to the perspective of awareness through the process of identification will, through their content, generate the impression of a personal self unbrokenly persisting from one conscious moment to the next. The stream of emotions and attendant thoughts also do much to explain the sense of a bounded self with *longer-term* identity. Their content tells the story of a continuing self who existed in the past and is planning for its future. That is because the current awareness identifies not merely with an isolated perspectival owner of the latest percept, thought or action, but with an integrated historical

personality who is the perspectival owner of the latest percept (etc.). Most emotions, moreover, have built into their content the idea of a numerically continuous self who has them. Having these emotions requires that one identify not only as their current personal owner, but also as the owner of a past or future scenario that was either perceived or is anticipated to trigger the emotion. For example, one could not feel guilty unless one believed that an earlier self – numerically identical to the current self – performed a regrettable action. One could not feel nervous at an impending speech unless one implicitly believed that a future self who will give the speech is identical to the current self who fears giving it. The presence of *taṇhā*, implicated in such emotions, also has the subject reflexively pre-supposing that it is an unbroken bounded self. It requires that one identify as the personal owner of past or future scenarios that are perceived, attentively or inattentively, as fulfilling or frustrating the desires upon which actual or imagined emotions depend.

In this boundedness tier, the self, viz., *unbroken*-awareness-*as-bounded*, comes out as *illusory* because unbroken awareness purports to be bounded, when in fact unbroken awareness is not bounded in its nature. Unbroken awareness is not bounded because (a) *unbroken awareness* in itself (at least from one conscious moment to the next) is unconstructed (from the arguments of Chapters 6 and 7) while (b) *unbroken*-awareness-*as-bounded* is constructed, viz., the content of an appearance contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since *unbroken* awareness is not constructed in actuality, it cannot be bounded (since unbroken-awareness-as-*bounded* is a construct) which renders the self, viz., *unbroken*-awareness-*as-bounded* an illusion.

With regard to unbroken awareness, the *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded unbroken awareness>, is presented as being *unconstructed*, in the manner of pure awareness. The self seems unconstructed, both in its moment-to-moment unbrokenness and in its longer-term identity. The *evidence* that we construe the self, viz., <bounded unbroken awareness>, as *unconstructed*, lies in the fact that the self and its boundaries never seem to exhibit any discontinuity, arising and passing away from the field of consciousness. If the self's unbrokenness (short- or long-term) *did* seem overtly constructed – dependent on impermanent, perspectivally ownable objects – then the self's 'unbrokenness' would seem more like *perdurance* (akin to Galen Strawson's 'pearl view' of the self, 1997, 424–425) with a decisively 'grainy' feel to its existence over time (perhaps, as mentioned in Chapter 4, like a string of mini general anaesthetics).

Our *explanation* for the impression of unconstructedness in the bounded and unbroken self (on a moment-to-moment basis) is that the process of identification will make it seem as if any appropriated thoughts are integral to awareness (hence unified and elusive), masking their actual coming and

going from awareness. If the appropriated thoughts *could* be directly observed to come and go – hence observed as discontinuous – then they would not *seem* integral to awareness in the first place. With the impression of longer-term identity, whenever current awareness identifies itself with a past or future perspectival owner (through memory or imagination), the native unbrokenness of current unconstructed awareness will make it seem as if this identified-as self – whether of distant past, imagined future or immediately preceding moment – is *itself* uninterruptedly persisting at every new moment of its imagined existence.

In this tier, the self, viz., *bounded-unbroken*-awareness-*as-unconstructed*, comes out as *illusory* because the *bounded-unbroken*-awareness purports to be unconstructed when as a matter of fact it is, as demonstrated in Chapter 8, a construct. The boundedness aspect of unbroken-awareness-as-unconstructed is the content of an appearance by which it is contributed to by perspectivally ownable objects that subtend *taṇhā*, in particular, thoughts and emotions. When we reflect in further detail, it is especially clear that the numerically bounded self with long-term personal identity is, due to its constructed status, not as it seems. It must have gaps in its existence since the identified-with ideas, even if continuously overlapping during waking hours, do not occur twenty-four hours a day. There is solid evidence that there are no thoughts during episodes of deep sleep. No thoughts imply no identification, and hence on our analysis, no constructed self. The 'self' that awakens from deep sleep cannot, therefore, be numerically identical to the self who fell asleep – contrary to how it would seem. Hence the self, viz., bounded-unbroken-awareness-*as-unconstructed*, is an illusion.

6. Invariability

With regard to invariability, the *first tier* of the self-illusion consists of input from the 'boundedness feeder' where the self is construed to be '*invariable*-awareness-*as-bounded*'. As with unbrokenness, I suggest that invariability, native to awareness from one conscious moment to the next, helps create the impression of a self that is invariable not only from moment to moment, but over the longer term. *Evidence* that we do indeed reflexively construe invariable awareness to be *bounded* lies in the fact that the quality of sameness that pervades conscious experience (from one moment to the next) is not merely impersonal but seems imbued with the essence of me, of *who I am*. The quality of 'me-ness' seems to persist over longer spans of time, despite any changes to my personality. No one else could possibly have this quality – it seems uniquely integral to my identity.

Our *explanation* for this *apparent boundedness* of invariable awareness is as follows. Feelings of personal ownership (explicit or implicit) towards various items trigger (or are triggered by) the reciprocal sense of being a personal

owner of that item, such that the process imparts, as it were, a 'tone' or 'hue' to the (otherwise colourless) subjective sense of presence. To use a crude metaphor, it is as if each instance of identification (of which there are many) spills a pod of dye into the colourless substrate of awareness, contributing to the resultant impression of a 'unique me'. The dye seems to 'colour' the awareness with the shade of the current (and often re-current) appropriated thoughts and ideas such that the subject feels integrated and defined in relation to that 'colour', sometimes more noticeably than usual (think of punkish persona versus owner of a breath). And like genuine dye, the mixing of many colours will result in the impression of a uniform colour. New colourful additions will soon swirl into this old colour – or near enough so that the change goes unnoticed. The uniform colour, then, is 'me', viz., the sense of being a *personal owner qua* one's general personality as it stands in relation to some item of the world (internal or external).

This content of the impression, <personal owner>, is the upshot of awareness identifying with a perspectival owner of some item, X (often X will be the sensory medium by which an external object is known to the subject). To each such identification, awareness brings the baggage of past identification: thoughts, memories, ideas of personality, etc., such that it seems like a perspectival owner *qua* integrated personality (etc.) who is the owner of X. The identification will have triggered new feelings of personal ownership towards X, implying the sense of *being* X's integrated personal owner – and this sense of being a personal owner of X will, in turn, feed back into the conception of personality. The resultant sense of a *self* is thus reactivated and revised upon each new identification on the part of awareness, with a perspectival owner (*qua* personality) of X. Because of the genuine overall resemblance between one identified-with 'personality-bite' and the next (along with the fact that each identification has awareness as invariable 'substrate') the subtle changes are likely to be glossed over in the manner that David Hume (1739, 169–170) describes. Even a sudden overlay of reds and purples (e.g., becoming a born-again Christian) will not fully disguise the underlying 'colour of me' (hence the sense that *I* have become a born-again Christian) and will sooner or later blend into the general 'dye-pool'.

The emotions and their attendant thoughts markedly contribute to the impression of it being the invariably same bounded me existing over time, since they elicit story lines of other times that we felt that way. For example, when feeling depressed, we are likely to remember the previous times we felt depressed, which will seem to create arcs of identity to the previous owners of the depression. Immediately there will seem to be a solid, numerically continuous and invariable *me* who is defined and hence bounded by the implicit quality of feeling depressed. (Which may explain why when someone jovially says 'that mood won't last' it is hard to believe!)

In this boundedness tier, the self, viz., *invariable*-awareness-*as-bounded* comes out as *illusory* because invariable awareness purports to be bounded, when in fact invariable awareness is not bounded in its nature. Invariable awareness is not bounded because (a) *invariable awareness* in itself is unconstructed (from the arguments of Chapters 6 and 7) while (b) *invariable*-awareness-*as-bounded* is constructed, viz., the content of an appearance contributed to by perspectivally ownable objects – in particular, thoughts and emotions subtending *taṇhā* (see explanation above). Since *invariable* awareness is not constructed in actuality, it cannot be bounded (since invariable-awareness-as-*bounded* is a construct) which renders the self, viz., *invariable*-awareness-*as-bounded*, an illusion.

With regard to invariable awareness, the *second tier* of the self-illusion consists of input from the 'awareness feeder' where the self, viz., <bounded invariable awareness>, is presented as being *unconstructed*, in the manner of pure awareness. The *evidence* that we construe the self, viz., <bounded invariable awareness> as *unconstructed*, lies in the fact that the underlying quality of me-ness seems to remain qualitatively the same either from one moment to the next or throughout my life, despite changes to personality. If the 'invariability' seemed overtly constructed by objects, then the 'I' would seem to change its essence to reflect the different identified-with objects at different times. This would effectively 'water down' any sense of invariability in a manner similar to how *gappiness* would 'water down' any sense of the self's unbroken persistence. The self does not, however, present itself (through its unbrokenness and invariability) as being 'watered down' in this fashion.

Our *explanation* for the impression of this unconstructed invariable awareness of a bounded self is that the native unconstructed awareness serves as the invariable and colourless 'substrate' into which the dyes of identity are spilt. Without this 'substrate' of awareness there could be no such sense of unbrokenness or invariability – short- or long-term. Just as spectral colours split by a prism share the deeper quality of luminosity, each moment of 'dyed awareness', although different in some way, will through a lifetime, share the deeper invariable quality of awareness. The mechanism of identification – with the help of factors alluded to in Tier One – somehow manages to juxtapose the genuine invariability of awareness onto 'dye-level', such that it feels as if the personal 'dye of self' is of an invariable hue. This impression may be enhanced by the natural 'change-blinding' effect that identification has upon the actual coming and going of identified-with ideas from a subject's awareness. So when we identify with a perspectival owner of X (creating the impression of being X's personal owner) there may be the tacit impression that we have *always been* the personal owner of X because the explicit introduction of that object into our sense of identity will have gone unnoticed.

In this 'unconstructedness' tier the self, viz., *bounded-invariable*-awareness-*as-unconstructed*, comes out as *illusory* because the *bounded-invariable*-awareness

purports to be unconstructed when as a matter of fact it is, as demonstrated in Chapter 8, a construct. The boundedness aspect of invariable-awareness-as-unconstructed is the content of an appearance by which it is contributed to by perspectivally ownable objects that subtend *taṇhā*, in particular, thoughts and emotions. The 'dye of self' is not invariable in the unconstructed manner of pure awareness. During waking hours, its hue is constantly and subtly changing. Hence the self, viz., bounded invariable-awareness-*as-unconstructed,* is an illusion.

7. Summary

The proposed model of the two-tiered self-illusion is by no means complete in its detail, although it does appeal to the arguments of previous chapters in its support. It offers an outline for a theory of self, which, following Buddhism and Advaita Vedanta, takes seriously the input of co-specified awareness as not constructed or illusory. By alluding to the process of identification (through which boundedness comes about), the account also builds a subject – with its psycho-physical perspective – into the overall construct of self.

Conclusion

I have drawn upon other chapters to provide an overarching account of how cognitive input could coalesce into the 'two-tiered illusion of self'. On such a model, the sense of the self's boundedness is kept alive by processes of identification, by which the subject assumes various items to be integrated or personally owned by itself. This both induces (and is induced by) a stream of *taṇhā*-driven thoughts and emotions whose content depicts a bounded self as the protagonist. In Chapter 8, I found strong support for this part of the theory from both the East (Buddhism) and the West (Antonio Damasio, 1999). The process of identification also brings to the self all the natural features of awareness, since it is a *subject with awareness* that goes about identifying with items. Identification serves to distort those natural features into the impression of a bounded *self* that harbours them. In this chapter, I tried to further describe how this might happen.

The multi-faceted role given to unconstructed awareness in the ontology of 'no-self' is what sets this proposed theory apart from many Western counterparts – theories proposed by Hume, James, Flanagan, Dennett and Damasio. The contribution to the self-illusion from unconstructed awareness is what notably aligns this project's theory of the self with those accounts that I have extracted from Buddhism (and also from Advaita Vedanta). It is these Eastern traditions to which I owe the fundamental ideas of this book, culminating in 'the two-tiered illusion of self': a model that I have tried to defend within a framework of Western analytic philosophy.

Glimpses Beyond

In this book I argued that the self is an illusion that draws its input from two tiers: one constructed (promoting boundedness) and one unconstructed (issuing from underlying awareness that is elusive, unified, unbroken and invariable). By not regarding unity and unbrokenness (in particular) as the primary source of the self's constructed and illusory status, the model of self-illusion proposed here – prefigured in Buddhism and Advaita Vedanta – significantly departs from the way the existence of self has been typically denied in the West. The departure can been seen as significant not only because the model is itself atypical, but because the model sets the stage for further enquiry in a way that other models of 'no-self' cannot accommodate. I talked about this enquiry in the Preface and will conclude this book by saying a few more words about it.

The model of a two-tiered self-illusion can be used as a foundation to begin investigating a question whose eventual answer could have great significance for the Western philosophy of mind, namely: is nibbāna psychologically possible? Framed specifically in terms of the two-tiered self-illusion, the question that would *begin* this enquiry could be as follows: is it psychologically possible that a person's awareness, with all its specifying features, could somehow divest itself of the self-illusion by directly 'waking up to' (i.e., *realising*) its real nature, just as one might wake up and realise that the ringing sound comes from the alarm clock and not the dreamt-of shrill voice? It has already been suggested, through such cases as epileptic automatism, that awareness can come apart from the sense of self. Damasio's patients, who seemed to temporarily lose the sense of themselves as bounded entities, did not cease witnessing; they remained awake and minimally responsive to their immediate surroundings. Yet there was not a shred of evidence that such patients – at least those Damasio described – had *any clue* as to what was happening. Stuck in the micro-context of the moment, there seemed to be no 'mental room' for them to harbour the realisation that an illusion of self was overthrown (and no 'mental room' for the usual identification). Many subjects of depersonalisation, on the other hand, seemed to retain their mental acuity (recall part of the DSM-IV definition: 'reality testing remains intact') and while evidently not identifying as the personal owner of various (perspectivally owned) thoughts and perceptions, they still harboured negative emotions about their general predicament. I argued that such emotions (through *taṇhā*) evidence a sense of personal ownership – and hence, a sense of self. And once again, these cases came across as being pathological, as impairing the well-being and normal functioning of the recipient.

The question worth pursuing, then, is whether awareness might 'wake up' from the illusion of self such that *contra* Damasio's patients, the separation of awareness from self-sense happens in a manner that is non-pathological and preserving of mental acuity and, *contra* typical depersonalisation, in a way that, of course, preserves no residual sense of self. Such a person (*qua* their awareness) will actively and suddenly *know*, with full mental acuity, that the illusion of self has been overthrown. This 'awakening' will involve the direct, clear and immediate realisation that awareness is not – and never was – intrinsically bounded in the way that the self is. It makes sense to suppose that, on a psychological level, such 'awakening' would be somewhat akin to a global depersonalisation – truly global – where *none* of one's perspectivally owned experiences, while clearly apprehended, are felt as being personally 'mine'. Unaccompanied by negative emotions, or by such impressions as being dead and 'falling deep within oneself', this mode of awareness – with no sense of personal 'me' or 'mine' whatsoever – would genuinely lack an accompanying sense of self.

The psychological possibility of awareness 'awakening' from the illusion of bounded self might be dismissed as fanciful speculation were it not for the fact that it is deemed, as we know, to be integral to the goal of practice in Buddhism (and also Advaita Vedanta). Nibbānic consciousness, we can recall from Chapter 2, is characterised in the Buddhist suttas as involving a mode of conscious awareness where there is no *taṇhā* or emotional investment in any item and hence, on our theory, no sense of self. It is notable that the occasional person reported to have 'realised' nibbānic consciousness (the *Arahant*) is often described as having 'woken up to their real nature'. This mode of knowing is reported to involve immense peace, happiness and mental clarity with complete freedom from the possibility of mental suffering or *dukkhā*. It is hence clearly depicted as being non-pathological.

In Chapter 1, I outlined the Four Noble Truths in an attempt to relate Buddhist principles to the concerns of this project. We may recall that the Fourth Noble Truth sets out a path of practice known as the 'Noble Eightfold Path' which divides into three main modes of practice called 'insight-wisdom' (*paññā*), 'meditation' (*samādhi*) and 'virtue' (*silā*). In light of the model of the two-tiered self-illusion, this system of training can be understood as one whose goal is to (a) dismantle the co-dependent sense of self and *taṇhā* while (b) simultaneously uncovering the real character of the native awareness such that it comes to reflexively know itself as it is in itself (as opposed to it mistakenly assuming it is a self).

In view of this, I see the natural line of enquiry as being one that addresses the following questions. Is it possible for a person's awareness to be divested of the self-illusion in such a manner that the person will suffer no pathological impairment, *contra* to what might be inferred from Damasio's theory? Could the methods of the Noble Eightfold Path, which purport to effect this non-pathological transformation, be amenable to philosophical scrutiny such

that their *prima facie* plausibility might at least be established? I am fully convinced that these are questions on which significant philosophical headway can be made without undue departure from the tradition of analytic philosophy. Having already given some thought to how the method of *meditation* could help undermine the sense of self in such a way that averts pathology, I present the following as a brief synopsis of my thinking on this subject so far.[1]

The general idea is that meditation would work, at least in part, by 'reprogramming' our usual patterns of attention so that the attention would no longer be compulsively captured and lost in the content of those 'story lines' needed to preserve the sense of a bounded self. I suggest, in line with Buddhism (and Damasio's findings),[2] that our ordinary patterns of attention are, to borrow a phrase from Hume, a 'slave to the passions', the passions being the stream of emotions and thoughts that help constitute *taṇhā* and the sense of a bounded self. While enslaved, the attention is repeatedly drawn into thoughts and story lines whose content implicitly depicts the self as protagonist of recalled or imagined scenarios in the past and future. In the broadest terms, I propose that Buddhist meditation would work via a system of training that gradually alters the patterns of attention so that they are no longer a 'slave to the passions'. If such methods enable the attention to become knowingly freed from the grip of such deep-rooted conditioning, then it would make sense to suppose that the quality of attention will, as a result, become very powerful and present-centred. If attention is no longer drawn compulsively and selectively to objects of perceived relevance to 'self', with its imagined time zones, then the *taṇhā* which relies upon this 'enslaved' pattern of attention will be destroyed. When *taṇhā* is destroyed, then so too, on the two-tiered model of self-illusion, will the sense of self-other boundedness. The non-bounded and unconstructed awareness, which is focused during attention – and which serves as the vehicle for meditation – will remain intact.

I suggest that a differential quality in attentive capacity could partially explain why a loss to the sense of self through methods of meditation would be non-pathological, while a loss to the sense of self through neurological trauma would be pathological. In all the pathological cases (such as epileptic automatism, Alzheimer's disease or akinetic mutism) Damasio has explicitly noted that the quality of attention is abnormally low. He has also noted that higher-quality attention is a reliable indicator of mental acuity (1999, 182–183). Because of this, the present-centredness of awareness through meditation would be just a distant cousin of that present-centredness forced upon patients undergoing an epileptic automatism. High-quality attention is not (to my knowledge) found in any recognised pathologies that fully suspend the sense of a self. While this factor will not be enough to explain how a person could autonomously survive in the absence of what Damasio calls 'over-riding individual concern', the calibre of attention purportedly

attained by those advanced in their meditation practice could well provide an important *clue* in understanding how pathology might be averted.

Within the framework of this project, a philosophical enquiry into how meditation might work (to undermine the sense of bounded self) would need to be supplemented by empirical research. Longitudinal studies may for instance be conducted on serious meditation practitioners to see whether their mode of practice does in fact undermine their sense of self. In Chapter 8, I argued, with support from Damasio's research, that there is a sense of self if and only if there is *taṇhā* – the investment of emotion in various desired situations such that there is mental *dukkhā* if the desires are frustrated. Damasio's account suggested that such 'boundary-preserving' emotions are neurophysiological, with distinct regions of the brain and body involved in their activation. I have already suggested that the presence or absence of *taṇhā* should in practice be empirically testable on several levels, from a neurophysiological level (e.g., those concerning what Damasio calls the 'proto-self') to a more overtly physical level (e.g., galvanic skin-conductance tests in response to actual and imagined scenarios).

If it turned out that the sense of self could indeed be non-pathologically lost through meditation practice, the implications for the philosophy of mind would be far-reaching – even if the resultant consciousness did not have all the alleged qualities of unconditioned *nibbānic consciousness*. For one thing, it would challenge the persuasive idea, forwarded by Damasio, that a sense of self is needed for well-functioning consciousness. Damasio writes:

> When the mental aspect of self is suspended, the advantages of consciousness soon disappear. Individual life regulation is no longer possible in a complex environment. In the full personal social sense, individuals remain capable of basic and immediate bodily maintenance. But their connection to the environment on which they depend is broken down, and, because of the breakdown, they cannot sustain such bodily maintenance. In fact, left to their own devices, death would ensue in a matter of hours because bodily maintenance would collapse. This, and comparable examples, would suggest that a state of consciousness which encompasses a sense of self as conceptualized in this book is indispensable for survival.
>
> (1999, 203–204)

Damasio's conjecture is based upon his observation of such cases as epileptic automatism, Alzheimer's disease and akinetic mutism – and as far as these cases go, it is a reasonable conjecture to make. He is not entirely unaware, however, that an unexplored avenue exists, since at one point he mentions, somewhat hesitantly, that 'impaired extended consciousness possibly contributes to the dissolution of self associated with states of depersonalization and with states of mystical [sic][3] selflessness' (Damasio, 1999, 216). Damasio

shows no indication of actually having *researched* those cases that could provide the best counterexamples to his idea that a sense of self is *always* needed for unimpaired consciousness. If there *could* be a different and non-pathological way for the sense of self to be lost, but with the consciousness remaining unimpaired, then this would compel the philosopher of mind to think about consciousness and its capacities in a new and different light. For instance, one would have to explain how a person could be motivated to act when there is *no* emotional investment in the results of any action, and hence, no overriding self-concern. On this matter, the Eastern accounts of how *Arahants* act in the world are telling. Not only are such persons reportedly able to 'get by'; their ordinary actions – those normally associated with autonomy and survival – are depicted, without exception, as proceeding more effortlessly and efficiently than the comparable actions from persons *with* a sense of self. From an ethical perspective, their conduct is invariably described as exemplary in its virtue, wisdom and compassion – exceeding, even, Aristotle's *phronimos*. Should these reports be correct, they would raise a plethora of questions, including the question of how it would be possible to feel compassion – and *dukkhā*-free happiness for that matter – in the absence of any 'self'-propelled emotions. To turn Damasio's conjecture on its head: could there be something about the ubiquitous sense of self and *taṇhā* that actually *prevents* consciousness from functioning at its best?

With these thoughts we leave the discussion.

Notes

Introduction

1. From John Bullitt (2005) 'What is Theravada Buddhism? A Thumbnail Sketch'. In http://www.accesstoinsight.org/theravada.html. This link may be helpful to readers who are unfamiliar with Buddhism (although a more detailed account of the central teachings is offered in this project).
2. The tradition of Advaita Vedanta is generally associated with the leading figure of Śaṃkara who lived in India around 788–820 CE. The teachings of Advaita Vedanta (that Śaṃkara and various sages such as Ramana Maharshi and Nisargadatta Maharaj re-interpreted) are based upon the Upaniṣads, some of which predate the Buddha. References to Advaita Vedanta will be relatively scarce in this project.

1. Some central distinctions and the Four Noble Truths

1. A published variation of this online paper appears in *Journal of Consciousness Studies* in 1997. Since the online version (which appears on a JCS discussion thread) is more suited to this chapter's discussion, it is the version I shall stick with. As no information is provided about its year of appearance, I shall assume the year to be 1997.
2. 'Qualia' is the plural term (singular: 'quale') that philosophers give to raw felt qualities of sensation such as the itchiness of hay fever or smell of fresh coffee or sadness of a memory.
3. The best one can hope for, if one is a physicalist, is some kind of identity, reduction or elimination of the witnessing perspective to brain-process, meaning that when the relevant brain-process is observed by a third party, then so is the witnessing perspective – should it exist. But note that the observed object would no longer be a subject *qua* subject, but merely (at the most) a subject *qua* brain-process. The subject–object distinction would not be obviously transgressed.
4. The division of conditioned co-dependency into explicit 'synchronic' and 'diachronic' dimensions is my own. However, both are implied in an abstract formulation of the doctrine, expressed in the suttas (and quoted in Harvey), hence 'That being, this comes to be; from the arising of that, this arises; that being absent, this is not; from the cessation of that, this ceases' (Peter Harvey, 1990, 54). On this, Harvey comments: 'This states the principle of conditionality, that all things, mental and physical, arise and exist due to the presence of certain conditions, and cease once their conditions are removed: nothing (except *Nibbāna*) is independent' (Harvey, 1990, 54).
5. See Armstrong, Martin and Place (1996) especially 71–87 and 184–191.
6. A good explication of these laws can be found in Bhikkhu Payutto's book *Good, Evil and Beyond: Kamma in the Buddha's Teachings* (1993, 1–2).
7. See for example Harvey (1990, 32) who writes, 'The Buddhist view, in fact, is that there is no known beginning to the cycle of rebirths and the world … However far back in time one goes, there must have been a prior cause for whatever beings existed at that time'.

8. Descriptions of the five *khandhās* can be found in most Buddhist texts, including Harvey (1990, 49–50) and Gethin (1998, 31–32).
9. The notion of 'real', in relation to the concept of 'illusory', will be properly analysed in Chapter 5.
10. This kind of belief in the self's existence is thus notably different from an intellectually motivated belief in the self's existence. With an intellectual belief (such as that which may be had on the basis of philosophical argument) the proposition believed is the content of a *thought* or *idea* that can be focused upon by the subject that appears to hold it. We must not confuse a thought or idea about X with a direct perception or conscious impression of X itself (compare the object of attention in a *thought* about a chair with that of a *perception* of a chair).
11. Suttas, drawn from the Pali Canon, are the closest we have to original teachings by the historical Buddha. Many of the Pali words have similar-sounding Sanskrit translations, for example, *nibbāna* (in Pali) is *nirvāna* (in Sanskrit); *kamma* (in Pali) is *karma* (in Sanskrit).
12. Adapted from a translation by Thanissaro Bhikkhu (2006). I translate the Pali term *dukkhā* as 'suffering' rather than 'stress'.
13. Readers wishing to gain another overview on the Four Noble Truths can consult Buddhist sources such as Gethin (1998, 59–84), Harvey (1990, 47–72) and Luang Por Sumedho (1992).
14. It may be pointed out that sometimes threads of pain or suffering can actually enhance the overall hedonic tone of an experience, as the pain of a mountain-climber might add to his overall elation at having climbed the mountain. This does not undermine our account of *dukkhā*, however, since Buddhism is not committed to *denying* that *dukkhā* can play this kind of role.
15. In the spectrum of ethical theories, Buddhism fits most closely with virtue ethics. The moral value of an action depends upon the state of mind (virtuous or vicious) from which the action springs. A very good book on *kamma* (with citations from various suttas) is Bhikkhu P.A. Payutto, *Good, Evil and Beyond: Kamma in the Buddha's Teaching*. Thailand: Buddhadhammā Foundation Publications, 1993.
16. So on the Buddhist position, rebirth does *not* involve some essential soul substance being reborn into another body; it rather involves the continuation of a causal process among the five *khandhās*. Such a process will 'carry' *kammic* imprints from one life to the next.
17. The Stoics seem a notable exception, however.
18. My interpretation of the concept of *nibbāna*, supported by citations from various suttas, will be elaborated in Chapter 2.
19. In Chapter 2, I provide evidence from the suttas to suggest that there may well be witnessing for the *Arahant* who dies; such witnessing would lack any object or external limitation by conditioned parameters.

2. Nibbāna

1. I nonetheless confess to preferring Thanissaro Bhikkhu's translations, hence, I rely mainly upon these in Chapter 3 where nothing too controversial hinges on them. The main controversy lies in my reading of the *nibbāna* suttas that I bring to the no-self doctrine – so it is important that the bulk of these are not translated by scholars, such as Thanissaro Bhikkhu, whose position I sympathise with. When I do refer to Thanissaro Bhikkhu's translations in this chapter, it is either uncontroversial or in a context where I appeal explicitly to a point that he has raised.

2. Since nothing too controversial hinges upon this sutta, I use Thanissaro Bhikkhu's (2005) interpretation here.
3. Translated by John Ireland (2006).
4. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
5. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999). A similar refrain occurs in AN IX. 36: 'Having viewed them thus, his mind then turns away from those states and focuses upon the deathless element: 'This is the peaceful, this is the sublime, namely, the stilling of formations, the relinquishment of all acquisitions, the destruction of craving, dispassion, cessation, *Nibbāna*.'' (*ibid*). In a later section, my analysis will imply that the term 'focusing', in this context (on *nibbāna*), cannot be taken literally if understood to mean 'paying attention'. 'Is percipient of' is more accurate.
6. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
7. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
8. Buddhist suttas do not specifically deploy these Kantian terms to describe the manner in which nibbānic consciousness is unconditioned; the description of nibbānic consciousness in these particular Kantian terms is my own.
9. Translated by Thanissaro Bhikkhu (1993a, 32).
10. Translated by Bhikkhu Bodhi (2000).
11. Translated by Bhikkhu Bodhi (2000).
12. Translated by Bhikkhu Bodhi (2000).
13. Translated by John Ireland (2006).
14. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
15. What if, for the sake of argument, all living beings were to be destroyed? Would there still be nibbānic consciousness? Given its timeless non-dependence upon any *khandhā*, mental or physical, the answer must be 'yes'. It is just that the mind of the *Arahant*, when freed from *taṇhā* and other 'covering' *khandhās*, is uniquely conducive to fully experiencing (and indeed, *being*) its intrinsic nature (nibbānic consciousness) as it is in itself.
16. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
17. *Nibbāna* would hence not be physical, since it would not depend in any way upon the body or brain. Why then, its association with properly functioning brains (as opposed to say tables or dead brains) or, in Buddhist terms, the *psycho*-physical *khandhās*, for example, those of an *Arahant*? Any role played by the *khandhās* must be understood in terms of their holding in the right relations to allow for the various degrees of revealing (rather than creating) the intrinsic nibbānic consciousness. The *khandhās* hence serve to cover, in various degrees, what is timelessly ever-present.
18. Translated by Thanissaro Bhikkhu (1993a, 28).
19. A detailed depiction of the *Arahant's* general mindset, supported by suttas, can be found in Thanissaro Bhikkhu (1993a, 95–120). Warpola H. Rāhula (1996, 43) also gives a nice depiction.
20. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).
21. Translated by Venerable Nyanaponika Therā and Bhikkhu Bodhi (1999).

3. The definition and status of self in Buddhism

1. It is reminded here that the indexical term 'itself' is not meant to convey the reality of a self.
2. Translated by Bhikkhu Nanamoli and Bhikkhu Bodhi (1995).
3. Translated by Thanissaro Bhikkhu (2004a).

4. Translated by Thanissaro Bhikkhu (2004b).
5. Translated by Thanissaro Bhikkhu (1993a, 75–76).
6. I say 'seemingly' in case it logically turns out that one person can access another person's experience, such as through clairvoyance. This would not impinge on our definition, since what matters is only that objects appear in a particular way from a subject's first-person perspective through virtue of occurring, as Philip Gerrans (2002, 37) puts it '*within her mind*, not the mind of someone else'. This manner of occurring is one that, as it happens, would *seem* to be accessed only by the subject to whom the object occurs. Perhaps, as Ronald de Sousa (2002) suggests, what I refer to here as 'perspectival ownership' can be divided further into 'titularity', 'privileged access' and 'incorrigibility'. If so, then it may be possible to treat 'perspectival ownership' as an umbrella term for these closely related subsets. The important matter, for our purposes, is that it broadly contrasts in the aforementioned way with the Buddhist notion of ownership.
7. Thus Tim Bayne (2004, 222–223) writes: 'Patients suffering from depersonalisation complain that their sensations and perceptions no longer feel as though they belong to them'. Later in this project, I will be discussing cases of depersonalisation that are outlined in psychological literature.
8. Bayne (2004) notes for instance that some authors have argued for a distinction between two kinds of ownership, namely, (1) 'a bare sense of being the *subject* of an experience' and (2) 'the sense of being its *author* or *agent*' (Bayne, 2004, 222). Bayne illustrates this with reference to a jingle running through one's head, where 'one experiences oneself as the subject of the auditory experience, but not its author' (*ibid*). Since one can clearly have a sense of my-ness without agency towards the jingle, Bayne wisely reserves the term 'ownership' for the notion of *subjecthood* as opposed to *agency* (in Chapter 4, I further justify this difference between ownership and agency). While improving on these authors, however, Bayne does not say *enough* about subjecthood to forge a clear distinction between perspectival and personal ownership. Later, it transpires that he must mean 'personal ownership' since he suggests that a subject can, through such experiences as depersonalisation, lose the sense of ownership towards their thoughts and experiences. His analysis fails to distinguish this from the legitimate sense in which such thoughts, etc., *will* remain (and will seem to remain) owned, namely, the perspectival sense.
9. In Chapter 4, when I say more about the concept of identification, it will be seen that the subject often identifies with an item through appropriating the *idea* of that item to their perspective.
10. Of course, as mentioned in Chapter 1, the subject will probably not be reflectively aware of taking on this identity; it will most likely be simply assumed as true in the relevant context.
11. The arguable exception, of course, is the consciousness *khandhā* that construes witnessing in its capacity of object-directedness. I argued in Chapter 2 that while this *khandhā* is conditioned in virtue of its object-oriented designation, the witnessing aspect to it is not conditioned.
12. Translated by Thanissaro Bhikkhu (1993a, 79–80). A similar theme is repeated in suttas throughout the *Khandhā Vaggā* (*Saṃs* XXII–XXXIV).
13. Thus, Hume (1740, 175) writes: 'But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness. I cannot discover any theory, which gives me satisfaction on this head'.
14. Translated by Thanissaro Bhikkhu (1993b). I insert the Pali words in lieu of his English translations.

15. See Miri Albahari (2002a). I have argued that many scholars, such as those cited in this section, conflate a common *reading* of Upaniṣadic teachings on *Atman*, a reading perhaps prevalent around the time of the Buddha, with the Upaniṣadic teachings themselves. This reading seems to construe ultimate reality on the 'ocean of sparkling consciousness' metaphor that I described in Chapter 2. When interpreted as Śaṃkara – or modern sages of the Advaita Vedanta tradition (such as Ramana Maharshi and Nisargadatta Maharaj) interpret them – it becomes much harder to defend the grounds upon which to separate the Upaniṣadic notion of *Atman* (at least, as they talk about it) from that of nibbānic witness-consciousness.
16. Be reminded that I use the term 'person' as I have defined it in Chapter 1 (a largely observable psycho-physical entity to whom selves are normally, even if mistakenly, ascribed) and that this notion of person is entirely relevant to the current discussion.
17. Some notable exceptions to this stance seem to include Peter Harvey (1995), Thanissaro Bhikkhu (1993a), (1993b), Christian Lindtner (1999), Jonathan Shear (1996), Arthur Deikman (1996), Robert Forman (1998) and Ken Wilber (2001). These authors do not depict Buddhism as construing a person's mind to be thoroughly impermanent – although they do not link nibbānic consciousness to *anattā* in the manner carried out here.
18. This succinct translation of the sutta was found at http://www.trinity.edu/rnadeau/Asian%20Religions/Lecture%20Notes/Mahayana%20Buddhism/Buddha%20Nagasena.htm. A longer version but with the same central terms including 'designation' and 'no permanent individual' is found in Bhikkhu Pesala (2001, 4–5).
19. We must note that the idea of nibbānic consciousness *intrinsically* being duped makes no sense, given that nibbānic consciousness is unconditioned by the four Kantian parameters. The *impression* of it being duped will be created by 'coverings' of the five *khandhā* types as affected by those aspects of nibbānic consciousness that 'show through'; hence, it will be the extrinsically conceived *pre*-nibbānic consciousness that is 'duped'.

4. The reflexively assumed self

1. The paper by Baron is unpublished and only available online (a fact that I verified by e-mail). The paper will be referred to again in a later section of this chapter.
2. Translated by Bhikkhu Nanamoli and Bhikkhu Bodhi (1995).
3. Nevertheless, I would argue that subjects of such depersonalisation episodes still appropriate a more general idea of 'me' to their perspective. It is just that none of their *perceptual* experiences during the episode seem to belong to this 'me'; more on this in Chapter 8.
4. I leave open the question as to whether reciprocal identification and personal ownership feelings are to be construed occurrently or dispositionally. For example, it might be enough that a subject who identifies with Y be disposed to feel personal ownership feelings in relation to her identification with Y. I suspect that the proposed analysis would be compatible with both readings.
5. For example, see Gilbert Ryle (1966), Chisholm (1969), G.E.M. Anscombe (1981, 21–36), and C.O. Evans (1970).
6. A recent formulation of the unity puzzle relevant to this project (and to be discussed later) is found in Bayne and Chalmers (2003).

5. How do we construe 'the self lacks reality'?

1. On this point, I have benefited from discussion with Jane McKessar.
2. I adopt the convention of referring to the <content of appearance> in these angle brackets, when the context renders it unsuitable to commit to one ontological status or another.
3. James makes it clear that the unity and unbrokenness we ascribe to the self is something unconstructed: '... common-sense insists that the unity of all the selves is not a mere appearance of similarity or continuity, ascertained after the fact. She is sure that it involves a real belonging to a real Owner, to a pure spiritual entity of some kind. Relation to this entity is what makes the self's constituents stick together as they do for thought' (James, 1890, 337).
4. This is not to deny that James' position differs from Hume's on the details of how the stream of thought preserves the impression of unity and identity. According to James (1890, 352), Hume did not properly account for the sense of sameness that one has over time.
5. Flanagan cites William James (1892), *Psychology: The Briefer Course*. G. Allport (ed). New York: Harper and Row, 1961.

6. Linking problems of consciousness with awareness

1. After making this distinction between 'tangible' and 'pure' awareness, Dainton (2002) goes on to discuss, and dismiss, the 'pure' variety (equivalent to a higher-order theory of consciousness). All the problems he points out seem due, precisely, to its non-tangible character. He does not discuss the 'tangible' variety.
2. Indeed, I think that it makes no sense to suppose, at least in the way that I have defined the terms, that anything *could* simultaneously be both subject and object.
3. Chisholm (1969, 95) thinks that 'both groups have lost their way' in supposing there to be a genuine problem here.
4. I am grateful to Hartley Slater for making this Rylean point to me so persuasively that it forced me to sharpen my formulation of the puzzle.
5. Any theory of consciousness that takes seriously (a) the apparent reality of phenomenal object properties (or *qualia*) and (b) analyses consciousness in terms of these properties, will satisfy this definition. It is usually this kind of consciousness that theoreticians have in mind when they address the 'hard problem' of consciousness. Chalmers (1996, 11) introduces a variant of this concept by listing a wide catalogue of sensory and perceptual experience, concluding the section by saying 'this brief look at the rich varieties of conscious experience should help focus attention on just what it is that is under discussion [in the hard problem]'. It should be noted, however, that he also includes in his list the 'deep and intangible' 'phenomenology of self', which he likens to a 'background hum' (Chalmers, 1996, 10), so it is not always clear whether he identifies the character of conscious experience exclusively with object-qualia. Were Chalmers to analyse consciousness in terms of *any* qualia, whether pertaining to object or subject, then the definition would become *too broad* to specifically substitute 'consciousness' in the elusiveness, unity and other puzzles to be discussed in this chapter.
6. A similar problem may arise in connection with Uriah Kriegel's (2004) account when he treats what he terms 'intransitive self consciousness' (viz., the phenomenology of self *qua* subject), as just another element of peripheral awareness. Kriegel's account seems to initially improve on Evans's account insofar as Kriegel

identifies only a select part of unprojected consciousness (rather than the whole of it) with the phenomenology of awareness, namely, with that aspect pertaining to one's peripheral *knowledge that* one is perceiving, thinking, etc. However, Kriegel does not think that this intransitive 'awareness of oneself', as he calls it, is fundamentally different (in kind) to peripheral awareness of other things (2004, 195), holding, even, that attentive awareness of ourselves as *subjects* is possible, although relatively rare (2004, 194). Kriegel's approach fails to capture the elusive '*modus operandi*' of 'subject' since it blatantly casts the so-called 'subject' as an *object*, that is, as something that can, in principle, be attended to.

7. Cited in Bayne and Chalmers (2003, 28). Bayne and Chalmers explore the notion of access consciousness for reasons that are similar to those outlined in this section, namely, to see whether the notion captures what seems unified about consciousness. Because of space limitations I do not include their reasons for regarding the notion of access consciousness as unsuitable for capturing the unity of consciousness.

7. The unconstructed reality of awareness

1. Imagine two worlds, W and W*, with individuals that are functionally and behaviourally identical. W has individuals with epiphenomenal conscious phenomena caused by neurological events; W* does not; everyone there is a zombie (with no inner conscious life). The only way to articulate the difference between W and W* is by appealing to the conscious epiphenomena, which would suggest that conscious phenomena have *prima facie* reality, *even if* they are causally impotent.
2. As said in a previous note, I do not know how to meaningfully construe the claim that something could present as both subject-like (to do with the *modus operandi* of witnessing) and object-like (as a witnessed object that can be attended to in principle).

8. How the self could be a construct

1. First-person accounts of depersonalization can be found on the website http://www.depersonalization.info/overview.html. Such people commonly report feeling a lack of emotion in response to triggers normally expected to cause emotion (e.g., a friend's death), and yet they invariably speak as if they long for their overall state to be overcome, sometimes even attempting suicide to escape their predicament.
2. 'BehaveNet - Clinical Capsule: DSM IV: Depersonalization Disorder' (2004) http://www.behavenet.com/capsules/disorders/depersdis.htm
3. 'AllPsych Online: The Virtual Psychology Classroom (Depersonalization Disorder)' (2004) http://allpsych.com/disorders/dissociative/depersonalization.html
4. 'Strangers to Our Selves' (2004) http://www.depersonalization.info/overview.html
5. For a website that has links to reputable articles on Cotard's syndrome, see http://groups.msn.com/TheAutismHomePage/cotardlinks.msnw
6. Other evidence for this is that early in his book, Damasio writes: 'In parallel with representing the printed words and displaying the conceptual knowledge required to understand what I wrote, your mind also displays something else, something sufficient to indicate, moment by moment, that *you* rather than anyone else are doing the reading and understanding of the text. … [This presence

signifies] you, as observer of the things imaged, owner of the things imaged, potential actor on the things imaged' (1999, 10). In Damasio (1999, 127), there is also clear indication that Damasio is concerned with personal ownership.

7. Perhaps in a similar fashion, a person with a phobia can develop a perceptual tropism, such that all items are perceived with an implicit 'filter' as to whether it is the feared object or not.
8. 'It is through feelings, which are inwardly directed and private, that emotions, which are outwardly directed and public, begin their impact on the mind ...' (Damasio, 1999, 36). See also Damasio (1999, 42).
9. This part of Damasio's theory (as with many other parts) is not merely speculative but has empirical support. It explains why, when the sense of self is impaired, the felt and observable emotions are also impaired, the same neurological machinery underpins them (Damasio, 1999, 100).
10. I inferred that epileptic automatisms involve awareness *sans* sense of self, from the fact that Damasio reported wakefulness *sans* sense of self in such patients – and wakefulness implies awareness. Awareness and wakefulness are not synonymous, however, since there is awareness during dream-sleep.
11. Although the word 'identification' was avoided so as to explain Damasio's theory in his own terms, identification is essentially what would, in his theory, bestow advantage to the organism in the transition between emotions merely being felt, and their being felt to belong to oneself. It would mark the transition between perspectival ownership of feelings (reacting instinctively to physical pleasure or pain) to a sense of personally owning them, such that they feel properly 'mine'. There would hence be identification with this perspectival owner of these feelings. While identification would herald the onset of that sense of personal boundedness that typifies the self, it would build on the existing tropism of a perspectival owner. As Damasio's theory makes plain, without this tropism, there could be nothing for the sense of self or identity to build upon. The perspectival owner is the first building block in the sense of self.
12. I say more about this in the concluding section entitled 'Glimpses Beyond'.

Glimpses Beyond

1. My paper (Albahari, 2002b), 'Can Heterophenomenology Ground a Complete Science of Consciousness' presents some of my further thoughts along these lines.
2. Antonio Damasio writes: 'To some degree, the message implied in the conscious state is: 'Focused attention must be paid to X.' Consciousness [with sense of self] results in enhanced wakefulness and focused attention, both of which improve image processing for certain contents and can thus help optimize immediate and planned responses (1999, 182). ... Emotion is critical for the appropriate direction of attention since it provides an automated signal about the organism's past experience with given objects and thus provides a basis for assigning or withholding attention relative to a given object (1999, 273). ... the consequences of having emotion and attention are entirely related to the fundamental business of managing life within the organism ...' (1999, 274). On Damasio's model, attention will quite naturally be directed by the emotions whose job it is to make sure – via the sense of self – that the bulk of focused-on objects are in some way relevant to the organism's welfare. For this reason, Damasio conjectures that the neurological mechanisms governing attention and those processing emotion and homeostatic balance will be in the same vicinity (1999, 274). What emerges,

then, is that the propensity for a person's awareness to selectively attend to things of perceived interest to the apparent self – integral to *taṇhā* viz., emotional investment – is necessary for maintaining the sense of self-other boundedness.

3. I must confess to cringing each time I see that meme-encrusted word 'mystical' being used to describe those 'selfless' modes of consciousness depicted in this project. For it does *nothing* to dispel connotations of such consciousness being so mysterious as to be entirely out of philosophy of mind's orbit and entirely unrelated to consciousness of the ordinary person!

Bibliography

Albahari, Miri. (2002a). "Against No-Atman Theories of Anattā" *Asian Philosophy* 12: 1. 5–20.

Albahari, Miri. (2002b). "Can Heterophenomenology Ground a Complete Science of Consciousness?" *Noetica: A Cognitive Science Forum*. http://www2.psy.uq.edu.au/CogPsych/Noetica/Articles/Noetica9.pdf

"AllPsych Online: The Virtual Psychology Classroom (Depersonalization Disorder)". (2004). http://allpsych.com/disorders/dissociative/depersonalization.html

Anscombe, G.E.M. (1981). *Metaphysics and the Philosophy of Mind (Vol 2)*. Oxford: Basil Blackwell.

Armstrong, David M. (1968). "Introspection" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.

Armstrong, David M., Martin, C.B. and Place, U.T. (1996). In Tim Crane (ed). *Dispositions: A Debate*. London: Routledge.

Austin, James H. (1999). *Zen and the Brain: Toward an Understanding of Meditation and Consciousness*. Cambridge, MA: MIT Press.

Austin, John. (1962). *Sense and Sensibilia*. Oxford: Oxford University Press.

The Autism Home Webpage. (2004). "Links to Additional Information on Cotard Syndrome and Related Conditions". http://groups.msn.com/TheAutismHomePage/cotardlinks.msnw

Bahm, Archie J. "Is There a Soul or No Soul? The Buddha Refused to Answer. Why?" In P.T. Raju (ed). *East-West Studies on the Problem of the Self*. The Hague: Martinus Nijhoff, 1968.

Bancroft, Anne. (1979). *Zen: Direct Pointing to Reality*. London: Thames & Hudson.

Baron, Richard. (2000). "The Self Is Unreal". http://www.rbphilo.com/

Bayne, Timothy (2004). "Self-Consciousness and the Unity of Consciousness" *The Monist* 87: 2. 224–241.

Bayne, Timothy and Chalmers, David J. (2003). "What Is the Unity of Consciousness?" In Cleeremans (ed). *The Unity of Consciousness*. Oxford: Oxford University Press, 23–58.

"BehaveNet – Clinical Capsule: DSM IV: Depersonalization Disorder". (2004). http://www.behavenet.com/capsules/disorders/depersdis.htm

Bodhi, Bhikkhu. (2000). *The Connected Discourses of the Buddha: A Translation of the Saṃyuttā Nikāya*. Boston: Wisdom Publications.

Bodhi, Bhikkhu. (1995). "Dhammā and Non-Duality - II" *Buddhist Publication Newsletter*. 29.

Bodhi, Bhikkhu. (1994). "Dhammā & Non-Duality - 1" *Buddhist Publication Society Newsletter*. 27.

Bodhi, Bhikkhu. (1990). *Nourishing the Roots: Essays on Buddhist Ethics*. Sri Lanka: The Wheel Publication.

Brahmavamso, Ajahn. (1999). *The Basic Method of Meditation*. Malaysia: Buddhist Gem Fellowship.

Brown, Jason W. (1999). "Microgenesis and Buddhism: The Concept of Momentariness" *Philosophy East and West* 49: 3. 261–277.

Bullitt, John. (2005). "What is Theravada Buddhism? A Thumbnail Sketch". http://www.accesstoinsight.org/theravada.html

Canfield, John V. (1990). *The Looking-Glass Self: An Examination of Self-Awareness*. New York: Praeger.

Cassim, Quassim. (1994). "Introduction" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.

Caston, Victor. (2002). "Aristotle on Consciousness" *Mind* 111: 444. 751–815.

Chah, Venerable Ajahn. (1992). *Food for the Heart*. Thailand: The Saṅgha, Wat Pah Nanachat.

Chah, Venerable Ajahn. (1979). *Clarity of Insight*. 2000. Thailand: Wat Pah Nanachat.

Chalmers, David J. (1999). "First-Person Methods in the Science of Consciousness" *Consciousness Bulletin*. http://consc.net/papers/firstperson.html

Chalmers, David J. (1996a). *The Conscious Mind*. New York: Oxford University Press.

Chalmers, David J. (1996b). "Facing up to the Problem of Consciousness" In Alfred W. Kaszniak, Stuart R. Hameroff and Alwyn C. Scott (ed). *Toward a Science of Consciousness: The First Tuscon Discussions and Debates*. 1996. Massachusetts: MIT Press.

Chisholm, Roderick M. (1969). "On the Observability of the Self" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.

Cohen, S.S. (1974). *Guru Ramana: Memories and Notes*. Madras, India: The Jupiter Press Private Ltd.

Collins, Steven. (1997). "A Buddhist Debate About the Self; and Remarks on Buddhism in the Work of Derek Parfit and Galen Strawson" *Journal of Indian Philosophy* 25. 467–493.

Conee, Earl. (1994). "Phenomenal Knowledge" *Australasian Journal of Philosophy* 72: 2. 136–150.

Dainton, Barry. (2002). "The Gaze of Consciousness" *Journal of Consciousness Studies* 9: 2. 31–48.

Damasio, Antonio. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. London: William Heinemann.

Damasio, Antonio. (1994). *Descartes' Error*. London: Macmillan General Books.

de Silva, Padmasiri. (1994). "Emotion Profiles: The Self and the Emotion of Pride" In Roger T. Ames, Wimal Dissanayake and Thomas P. Kasulis (ed). *Self as Person in Asian Theory and Practice*. Albany: SUNY Press, 1994.

de Sousa, Ronald. (2003). "Emotion" *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/emotion/

de Sousa, Ronald. (2002). "Twelve Varieties of Subjectivity: Dividing in Hopes of Conquest" In J. Lazarrabal and L.A. Perez Miranda (ed). *Knowledge, Language and Representation*. Dordrecht: Kluwer, 135–151.

Deikman, Arthur J. (2000). "A Functional Approach to Mysticism" *Journal of Consciousness Studies* 7: 11–12. 75–91.

Deikman, Arthur J. (1996). "'I' = Awareness" *Journal of Consciousness Studies* 3: 4. 350–356.

Dennett, Daniel C. (2001a). "Are We Explaining Consciousness Yet?" *Cognition* 79: 1–2. http://ase.tufts.edu/cogstud/papers/cognition.fin.htm. 221–237.

Dennett, Daniel C. (2001b). "The Fantasy of First-Person Science". http://ase.tufts.edu/cogstud/papers/chalmersdeb3dft.htm

Dennett, Daniel C. (1996). "Commentary on Chalmers: Facing Backwards on the Problem of Consciousness" *Journal of Consciousness Studies* 3: 1. http://ase.tufts.edu/cogstud/papers/chalmers.htm. 4–6.

Dennett, Daniel C. (1991). *Consciousness Explained*. London: Penguin Books.

Depersonalization.info. (2004). "Strangers to our Selves". http://www.depersonalization.info/overview.html

Depersonalization Research Unit. (2006). “About Depersonalization (What Is)” Kings College of London. http://www.iop.kcl.ac.uk/iopweb/departments/home/default.aspx?locator=911&context=718

Depersonalization Research Unit. (2001). “Depersonalization” *The Pyschologist* 14: 3. 128–133.

Descartes, René. (1641). *The Meditations*. Translated by John Veitch, U.S.: The Open Court Library of Philosophy, 1968.

Deutsch, Eliot. (1969). *Advaita Vedanta: A Philosophical Reconstruction*. Honolulu: University of Hawaii Press.

Dretske, Fred. (1999). “The Mind’s Awareness of Itself” *Philosophical Studies* 95. 103–124.

Dreyfus, Hubert L. (1993). “Review of Varela, Francisco J., Thompson, Evan, Rosch, Eleanor. *The Embodied Mind, Cognitive Science and Human Experience*“ *Mind* 102: 407. 542–546.

Edey, Mait. (1997). “Subject and Object” *Journal of Consciousness Studies* 4: 5/6. 526–531. Online version (slightly different) at http://www.imprint.co.uk/online/ edey.html

Evans, C.O. (1970). *The Subject of Consciousness*. London: George Allen & Unwin Ltd. http://www.mentalstates.net/SOC.html

Evans, Gareth. (1982). “Self-Identification” In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.

Flanagan, Owen. (2003) “The Colour of Happiness” *New Scientist*. 24 May 2003. 44.

Flanagan, Owen. (1992). *Consciousness Reconsidered*. Cambridge, MA: The MIT Press.

Forman, Robert K.C. (1998). “What Does Mysticism Have to Teach Us About Consciousness?” *Journal of Consciousness Studies* 5: 2. 185–201.

Garfield, Jay and Priest, Graham. (2003). “Nagarjuna and the Limits of Thought” *Philosophy East and West* 53: 1. 1–21.

Geertz, Clifford. (1975). *The Interpretation of Cultures: Selected Essays*. London: Hutchinson.

Gerrans, Philip. (2002). “Authorship and Ownership of Thoughts” *Philosophy, Psychiatry, and Psychology* 8: 2–3. 231–237.

Gethin, Rupert. (1998). *The Foundations of Buddhism*. Oxford: Oxford University Press.

Giles, James. (1993). “The No-Self Theory: Hume, Buddhism, and Personal Identity” *Philosophy East and West*. 43: 175–200.

Goleman, Daniel. (2003). *Destructive Emotions: How Can We Overcome Them? A Scientific Dialogue with the Dalai Lama*. New York: Bantam Books.

Goleman, Daniel. (1996). *The Meditative Mind: The Varieties of Meditative Experience*. New York: J.P. Tarcher.

Gruzalski Bart. (1996). “The Possibility of Nonattachment” In Ramakrishna Puligandla and David Lee Miller (ed). *Buddhism and the Emerging World Civilization: Essays in Honor of Nolan Pliny Jacobson*. Carbondale and Edwardsville: Southern Illinois University Press.

Gunaratana, Venerable Mahathera. (1991). *Mindfulness in Plain English*. Malaysia: Bhāvanā Society.

Guzeldere, Guven. (1996). “Consciousness and the Introspective Link Principle” In Alfred W. Kaszniak, Stuart R. Hameroff and Alwyn C. Scott (ed). *Towards a Science of Consciousness: The First Tuscon Discussions and Debates*. Massachusetts: MIT Press.

Hanh, Thich Nhat. (1995). *Zen Keys*. New York: Doubleday.

Harvey, Peter. (1995). *The Selfless Mind*. Surrey: Curzon Press.

Harvey, Peter. (1993). “The Mind-Body Relationship in Pali Buddhism: A Philosophical Investigation” *Asian Philosophy* 3: 1. 29–41.

Harvey, Peter. (1990). *An Introduction to Buddhism*. Cambridge: Cambridge University Press.

Heffner, Gary J. (2004) Cotard's Syndrome. The Autism Home Page. http://groups.msn.com/TheAutismHomePage/cotardsyndrome.msnw

Hoffman, Frank J. (1982). "The Buddhist Empiricism Thesis" *Religious Studies* 18. 151–158.

Hume, David. (1739). "Of Personal Identity" In John Perry (ed). *Personal Identity*. 1975. Berkeley: University of California Press.

Hume, David. (1740). "Second Thoughts" In John Perry (ed). *Personal Identity*. 1975. Berkeley: University of California Press.

Ireland, John D. (2006) (trans). "The Group of the Twos" *Iti II.16; Iti 37*. http://www.accesstoinsight.org/tipitaka/kn/iti/iti.2.042-049x.irel.html#iti-043.

Ishigami-Iagolnitzer, Mitchiko. (1997). "The Self and the Person as Treated in Some Buddhist Texts" *Asian Philosophy* 7: 1. 37–45.

Jackson, Frank. (1986). "What Mary Didn't Know" *The Journal of Philosophy*. 291–295.

James, William. (1890). *The Principles of Psychology: Vol 1*, Chapter X. From Internet resource "Classics in the History of Psychology" developed by Daniel Green: http://psychclassics.yorku.ca/James/Principles/prin10.htm

Janzen, Greg. (forthcoming) "Phenomenal Character as Implicit Self-Awareness" *Journal of Consciousness Studies*.

Kalupahana, David J. (1976). *Buddhist Philosophy: A Historical Analysis*. Hawaii: University of Hawaii Press.

Kant, Immanuel. (1787). *Critique of Pure Reason*. Translated by Norman Kemp Smith (1929) Palgrave Macmillan. E-text available at http://www.hkbu.edu.hk/~ppp/cpr/toc.html

Kapleau, Philip. (1965). *The Three Pillars of Zen*. Boston: Beacon Press.

Kasulis, Thomas P. (1997). "The Buddhist Concept of Self" In Eliot Deutsche and Ron Bontekoe (ed). *A Companion to World Philosophies*. Malden, MA: Blackwell Publishers.

Khemā, Ayyā. (1987). *Being Nobody, Going Nowhere: Meditations on the Buddhist Path*. Boston: Wisdom Publications.

Khemā, Ayyā. (1984). "Meditating on No-Self" In *Boddhi Leaves*. Sri Lanka: Buddhist Publication Society, No. 95.

Kitcher, Patricia. (2000). "On Interpreting Kant's Thinker as Wittgenstein's 'I'" *Philosophy and Phenomenological Research* LXI: 1. 33–63.

Kornblith, Hilary. (1998). "What Is It Like to Be Me?" *Australasian Journal of Philosophy* 76: 1. 48–60.

Kriegel, Uriah. (2005). "Naturalizing Subjective Character" *Philosophy and Phenomenological Research*, 71. 23–57.

Kriegel, Uriah. (2004). "Consciousness and Self-Consciousness" *The Monist* 87. 185–209.

Kriegel, Uriah. (2003). "Intrinsic Theory and the Content of Inner Awareness" *Journal of Mind and Behavior* 24. 171–198.

Lama, Dalai. (2003). "On the Luminosity of Being" *New Scientist* 42–43.

Larson, Gerald, J. (1988). "Introduction: The "Age-Old Distinction between the Same and the Other"" In Gerald J. Larson and Eliot Deutsch (ed). *Interpreting across Boundaries: New Essays in Comparative Philosophy*. 1988. Princeton: Princeton University Press.

Laycock, Steven W. (1998). "Consciousness It/Self" *Journal of Consciousness Studies* 5: 2. 141–152.

Lesser, A.H. (1979). "Eastern and Western Empiricism and the 'No-Self' Theory" *Religious Studies*. 15: 55–64.

Levine, Michael. (1998). "No-Self, Real Self, Ignorance and Self-Deception: Does Self-Deception Require a Self?" *Asian Philosophy* 8: 2. 103–110.

Lindtner, Christian. (1999). "From Brahmanism to Buddhism" *Asian Philosophy*. 9: 5–38.

Livingston, Paul. (2002). "Experience and Structure" *Journal of Consciousness Studies* 9: 3. 15–33.

Locke, John. (1690). *An Essay Concerning Human Understanding*. A.D. Woozley (ed). 1977. Glasgow: Collins Fount Paperbacks.

Lowe, E.J. (1997). "Self, Agency and Mental Causation". http://www.philosophy.stir.ac.uk/old/cnw/webpapers/lowe1.htm

Lyons, William E. (1986). *The Disappearance of Introspection*. Massachusetts: MIT Press.

Marks, Joel and Ames, Roger T. (ed). (1995). *Emotions in Asian Thought: A Dialogue in Comparative Philosophy*. New York: SUNY Press.

Mascaró, Juan. (trans). *The Dhammapada: The Path of Perfection*. (1973). London: Penguin Books.

Matilal, Bimal Krishna. (1994). "The Perception of Self in Indian Tradition" In Roger T. Ames, Wimal Dissanayake and Thomas P. Kasulis (ed). *Self as Person in Asian Theory and Practice*. New York: State University of New York Press.

McKinnon, Neil. (2003). "Presentism and Consciousness" *Australasian Journal of Philosophy* 81: 3. 305–323.

Nagel, Bruno. M.J. (1995). "Unity and Contradiction: Some Arguments in Utpaladeva and Abhinavagupta for the Evidence of the Self as Sivā" *Philosophy East and West* 45: 4. 501–525.

Nagel, Thomas. (1986). *The View from Nowhere*. Oxford: Oxford University Press.

Nanamoli, Bhikkhu and Bodhi, Bhikkhu. (1995) (trans). *Middle Length Discourses of the Buddha: A Translation of the Majjhimā Nikāya*, Somerville: Wisdom Publications.

Nanasampanno, Venerable Ācariyā Mahā Boowa. (1987). *Straight from the Heart: Thirteen Talks on the Practice of Meditation*. Thailand: Wat Pa Baan Taad.

Nanasampanno, Venerable Phra Acharn Mahā Boowa. (1982). *The Venerable Phra Acharn Mun Bhuridatta Therā: Meditation Master*. Thailand: WAVE Publications.

Nemirow, Laurence. (1990). "Physicalism and the Cognitive Role of Acquaintance" In W.G. Lycan (ed). *Mind and Cognition*. Oxford: Blackwell Publishers.

Nida-Rumelin. (1998). "On Belief About Experiences. An Epistemological Distinction Applied to the Knowledge Argument against Physicalism" *Philosophy and Phenomenological Research* LVIII: 1.

Nikolinakos, Drakon Derek. (2004). "Anosognosia and the Unity of Consciousness" *Philosophical Studies* 119. 315–342.

Nyanaponika Therā, Venerable. (1998). *Abhidhamma Studies: Buddhist Explorations of Consciousness and Time*. Boston: Wisdom Publications.

Nyanaponika Therā, Venerable. (1996). *The Power of Mindfulness*. Kandy: Buddhist Publication Society.

Nyanaponika Therā, Venerable and Bodhi, Bhikkhu. (1999) (trans & ed). *Numerical Discourses of the Buddha: An Anthology of Suttas from the Aṅguttara Nikāya*. Kandy: Buddhist Publication Society.

Nyanatiloka, Mahathera. (1994). *Fundamentals of Buddhism: Four Lectures*. Kandy: Buddhist Publication Society.

O'Brien, Gerard and Opie, Jon. (1998). "The Disunity of Consciousness" *Australasian Journal of Philosophy* 76: 3. 378–395.

Olsson, Erik J. (1997). "Coherence and the Modularity of Mind" *Australasian Journal of Philosophy* 75: 3. 404–411.

Osborne, Arthur (ed). (1971). *The Teachings of Bhagavan Sri Ramana Maharshi in His Own Words*. London: Rider and Company.
Osbourne, Arthur. (1954). *Ramana Maharshi and the Path of Self-Knowledge*. London: Rider and Company.
Panikkar, Raimundo. (1988). "What Is Comparative Philosophy Comparing?" In Gerald J. Larson and Eliot Deutsch (ed). *Interpreting across Boundaries: New Essays in Comparative Philosophy*. 1988. Princeton: Princeton University Press.
Parfit, Derek. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
Payutto, Bhikkhu P.A. (1993). *Good, Evil and Beyond*: *Kamma in the Buddha's Teaching*. Bangkok, Thailand: Buddhadhammā Foundation.
Perry, John. (1997). "Myself and I" In Marcelo Stamm (ed). *Philosophie in Synthetisher Absicht*. Stuttgart: Klett-Cotta, 1998, 83–103. http://www-csli.stanford.edu/~john/PHILPAPERS/myself.pdf
Perry, John. (1979). "The Problem of the Essential Indexical" *Nous*. 13(1): 3–21.
Pesala, Bhikkhu. (2001). *The Debate of King Milinda*. Penang: Inward Path.
Petranker, Jack. (2001). "Who Will Be the Scientists: A Review of B. Alan Wallace's 'the Taboo of Subjectivity'" *Journal of Consciousness Studies* 8: 11. 83–90.
Pickering, John (ed). (1997). *The Authority of Experience: Essays on Buddhism and Psychology*. Surrey: Curzon Press.
Powell, Robert (ed). (1994). *The Ultimate Medicine as Prescribed by Nisargadatta Maharaj*. California: Blue Dove Press.
Prasad, Hari Shankar. (2000). "Dreamless Sleep and Soul: A Controversy between Vedanta and Buddhism" *Asian Philosophy* 10: 1. 61–73.
Price, H.H. and Braithwaite, R.B. (1964). "Half-Belief" *Aristotelian Society* Suppl. 38. 149–162.
Rāhula, Walpola, H. (1996). *What the Buddha Taught*. Dehiwali, Sri Lanka: Buddhist Cultural Centre.
Roessler, Johannes. (2000). "Attention and the Self: An Appreciation of C.O. Evans' 'the Subject of Consciousness'" *Journal of Consciousness Studies* 7: 5. 76–82.
Ryle, Gilbert. (1966). "Self-Knowledge" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.
Śaṃkara. (788–820 CE). *Crest-Jewel of Wisdom*. Translated by John Richards. (from 1998). http://www.realization.org/page/namedoc0/vc/vc_0.htm
Sharma, Arvind. (1999). "Jivanmukti in Neo-Hinduism: The Case of Ramana Maharsi" *Asian Philosophy* 9: 2. 93–105.
Shear, Jonathan. (1996). "On the Existence of a Culture Independent Core Component of Self" In Ninian Smart and B. Srinivasa Murthy (ed). *East-West Encounters in Philosophy and Religion*. Long Beach, California: Long Beach Publications.
Shear, Jonathan and Jevning, Ron. (1999). "Pure Consciousness: Scientific Exploration of Meditation Techniques" *Journal of Consciousness Studies*. 6(2–3): 189–210.
Shoemaker, Sydney. (1986). "Introspection and the Self" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.
Shoemaker, Sydney. (1968). "Self-Reference and Self-Awareness" In Quassim Cassam (ed). *Self-Knowledge*. 1994. Oxford: Oxford University Press.
'Shravak', Lalji. (1999). "Buddha's Rejection of the Brahmanical Notion of Atman" *Communication and Cognition* 32: 1/2. 9–20.
Siderits, Mark. (1997). "Buddhist Reductionism" *Philosophy East and West* 47: 4. 455–478.
Siderits, Mark. (1987). "Beyond Compatibilism: A Buddhist Approach to Freedom and Determinism" *American Philosophical Quarterly* 24: 2. 149–159.
Smart, Ninian. (1999). "A Survey of Buddhist Thought" In Eliot Deutsch and R. Bontelcoe (ed). *A Companion to World Philosophies*. Oxford: Blackwell Publishers, 78–98.

Smart, Ninian. (1988). "The Analogy of Meaning and the Tasks of Comparative Philosophy" In Gerald J. Larson and Eliot Deutsch (ed). *Interpreting across Boundaries: New Essays in Comparative Philosophy*. Princeton: Princeton University Press.

Smith, David Woodruff. (1986). "The Structure of (Self-) Consciousness" *Topoi* 5. 149–156.

Strawson, Galen. (1997). "The Self" *Journal of Consciousness Studies* 4: 5/6. 405–528.

Strawson, Peter F. (1959). *Individuals*. London: Methuen & Co Ltd.

Stubenberg Leopold. (1996). "The Place of Qualia in the World of Science" In Alfred W. Kaszniak, Stuart R. Hameroff and Alwyn C. Scott (ed). *Toward a Science of Consciousness: The First Tuscon Discussions and Debates*. Massachusetts: MIT Press.

Sturgeon, Scott. (1994). "The Epistemic View of Subjectivity" *The Journal of Philosophy* XCI: 5. 221–235.

Sumedho, Venerable Ajahn. (1992). *The Four Noble Truths*. Hertfordshire: Amaravati Publications.

Thanissaro, Bhikkhu. (2006) (trans). "Dhammacakkappavattana Sutta" *SN LVI.11* http://www.accesstoinsight.org/canon/sutta/saṃyuttā/sn-56-011-tb0.html

Thanissaro, Bhikkhu. (2005) (trans). "Koṭṭhita Sutta" *AN IV. 174* http://www.accesstoinsight.org/canon/sutta/aṅguttara/an04-174.html

Thanissaro, Bhikkhu. (2004a) (trans). "Gadulla Sutta: The Leash" *SN XXII.99*. http://www.accesstoinsight.org/canon/sutta/saṃyuttā/sn22-099.html

Thanissaro, Bhikkhu. (2004b) (trans). "Vinā Sutta: The Lute" *SN XXXV. 205* http://www.accesstoinsight.org/canon/sutta/saṃyuttā/sn35-205.html

Thanissaro, Bhikkhu. (1993a). *The Mind Like Fire Unbound*. Barre, MA: Dhammā Dānā Publications.

Thanissaro, Bhikkhu. (1993b). "The Not-Self Strategy". http://www.accesstoinsight.org/lib/modern/thanissaro/notself.html

Tye, Michael. (1997). "The Problem of Simple Minds: Is There Anything It Is Like to Be a Honey Bee?" *Philosophical Studies* 88. 289–317.

Tzu, Lao. (1963). *Tao Te Ching*. Middlesex: Penguin.

Van Gulick, Robert. (2001). "Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A Philosophic Overview" *Journal of Consciousness Studies* 8: 9–10. 1–34.

Varela, Francisco J. and Shear, Jonathan. (ed). (1999a). *The View from Within: First Person Approaches to the Study of Consciousness*. Thorverton: Imprint Academic.

Varela, Francisco J. and Shear, Jonathan. (1999b). "First-Person Methodologies: What, Why, How?" *Journal of Consciousness Studies*.

Varela, Franscisco J., Thompson, Evan and Rosch, Eleanor. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.

Velleman, J. David. (2002). "Identification and Identity" In Sarah Buss (ed). *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, MA: Bradford-Book-MIT-Pr.

Velmans, Max. (1996). "Defining Consciousness". http://cogprints.ecs.soton.ac.uk/archive/00000395/00/Definingconsciousness.html

Von Glasenapp, Helmuth. (1978). *Vedanta and Buddhism: A Comparative Study*. Kandy, Sri Lanka: Buddhist Publication Society. http://www.accesstoinsight.org/lib/bps/wheels/wheel002.html

Wallace, B. Alan. (1999). "The Buddhist Tradition of Samathā: Methods for Refining and Examining Consciousness" *Journal of Consciousness Studies* 6: 2–3. 175–187.

Walsh, Roger. (2000). *The Taboo of Subjectivity: Toward a New Science of Consciousness*. Oxford: Oxford University Press.

Walsh, Roger. (1992). “Can Western Philosophers Understand Asian Philosophies?” In James Olgivy (ed). *Revisioning Philosophy*. Albany: SUNY Press.
Walshe, Maurice. (1995). *The Long Discourses of the Buddha: A Translation of the Dīgha Nikāya*. Boston: Wisdom Publications.
Watt, Douglas F. (2000). “Emotion and Consciousness: Part II: A Review of Antonio Damasio’s ‘the Feeling of What Happens: Body and Emotion in the Making of Consciousness’” *Journal of Consciousness Studies* 7: 3. 72–84.
Werner, Karel. (1996). “Indian Conceptions of Human Personality” *Asian Philosophy* 6: 2. 93–107.
Wilber, Ken. (2001). *The Eye of Spirit: An Integral Vision for a World Gone Slightly Mad.* 3rd edition. Boston: Shambala.

Index